

Département de géomatique appliquée  
Faculté des lettres et sciences humaines  
Université de Sherbrooke

Méthode d'obtention et de correction atmosphérique semi-automatisée d'imagerie Sentinel-2, dans  
un cadre de cartographie des sols agricoles en Montérégie

Benjamin Mainguy-Seers

Essai présenté pour l'obtention du grade de maître ès sciences géographiques (m.sc.), cheminement  
de type cours en géodéveloppement durable

Mai 2019

## TABLE DES MATIÈRES

Remerciements .....	1
résumé.....	2
1. Introduction.....	2
1.1. Contexte .....	2
1.2. Problématique .....	4
1.3. Objectifs .....	5
2. Cadre théorique.....	6
2.1. Granulométrie des sols .....	6
2.2. Télédétection et cartographie des sols.....	7
2.3. Services web d'imagerie .....	8
2.4. Correction atmosphérique .....	8
3. Matériel et méthodes.....	11
3.1. Site d'étude .....	12
3.2. Données .....	13
3.2.1. Imagerie .....	13
3.2.2. Points d'échantillonnage de sol .....	15
3.3. Méthodologie .....	19
3.3.1. Acquisition semi-automatique des images sentinel-2.....	19
3.3.2. Correction atmosphérique semi-automatique des images .....	20
3.3.3. Classification .....	23
4. Résultats .....	27
4.1. Cartographie .....	27
4.2. Statistiques des classifications .....	28
5. Discussion.....	32
5.1. Disposition géographique.....	32
5.2. Statistiques .....	33
5.3. Comparaison avec d'autres données et un autre classificateur .....	35
6. Conclusion .....	37
Bibliographie.....	39
Annexe.....	43

## LISTE DES TABLEAUX

Tableau 1. Indices spectraux pertinents . . . . .	7
Tableau 2. Granules Sentinel-2 utilisées . . . . .	14
Tableau 3. Points de granulométrie in situ . . . . .	16
Tableau 4. Nombre final de points par couverture . . . . .	18
Tableau 5. Distance de Bhattacharyya avec l'entièreté des classes (sauf organique) avec l'image #9 de l'étude . . . . .	23
Tableau 6. Distance de Bhattacharyya avec les classes suffisamment représentées (sauf organique) avec l'image #9 de l'étude . . . . .	24
Tableau 7. Divergence transformée avec les classes suffisamment représentées (sauf organique) avec l'image #9 de l'étude . . . . .	24
Tableau 8. Agrégation des classes en argile, loam et sable . . . . .	25
Tableau 9. Distance de Bhattacharyya et divergence transformée avec classes agrégées (image #9) . . . . .	25
Tableau 10. Disposition des classes . . . . .	28
Tableau 11. Statistiques principales des matrices de confusion . . . . .	29
Tableau 12. Matrices de confusion (%) après classification par maximum de vraisemblance pour l'ensemble des couvertures agrégées et l'agrégat des meilleures classifications . . . . .	29
Tableau 13. Matrices de confusion (%) après classifications par maximum de vraisemblance (par année de couverture) . . . . .	30
Tableau 14. Distances de Bhattacharyya (DB) et divergence transformée (DT) pour chaque couverture et les agrégats . . . . .	31
Tableau 15. Signification du coefficient de Kappa . . . . .	31
Tableau 16. Classes conservées pour la classification à deux classes (16/05/2018 . . . . .	35
Tableau 17. Matrice de confusion de la classification à deux classes (16/05/2018) . . . . .	35
Tableau 18. Résultats d'identification des textures de sol à partir de 12 images Landsat par la méthode AFD (analyse factorielle discriminante), étude de l'équipe d'OT d'Effigis . . . . .	36
Tableau 19. Images Sentinel-2 utilisées . . . . .	43
Tableau 20. Bandes captées par le MSI de Sentinel-2 . . . . .	44

## LISTE DES FIGURES

Figure 1. Triangle des textures de sol (version canadienne)	6
Figure 2. Trajet de la lumière captée par satellite	9
Figure 3. Étapes principales du présent travail	11
Figure 4. Zone d'étude	12
Figure 5. Représentation visuelle des indices spectraux avec zoom aléatoire	15
Figure 6. Points d'échantillonnage in situ du sol (agrandissement aléatoire sur la zone d'étude)	17
Figure 7. Zone d'étude selon la distribution des points et le nombre d'images	18
Figure 8. Affichage des images après exécution du script de sélection sur <i>Google Earth Engine</i> (Capture d'écran)	19
Figure 9. Fichier .txt de paramètres 6S pour une bande d'une image	21
Figure 10. Agrandissement sur l'image #4 (2018/054/23), avant et après correction atmosphérique	22
Figure 11. Méthodes de classifications sur PCI	26
Figure 12. Granulométrie de la zone d'étude, classifiée avec l'image #9, du 2018/05/21	27
Figure 13. Agrandissement sur le Mont Rougemont, autour duquel le loam est souvent « repéré »	32
Figure 14. Agrandissement sur l'image Sentinel-2 #1 (voir Tableau 2, section 3.2.1.)	49
Figure 15. Granulométrie de la zone d'étude, classifiée avec l'image #9, du 2018/06/10	50
Figure 16. Granulométrie de la zone d'étude, classifiée avec l'image #10, du 2018/05/31	51
Figure 17. Granulométrie de la zone d'étude, classifiée avec l'image #9, du 2018/05/21	52
Figure 18. Granulométrie de la zone d'étude, classifiée avec l'image #8, du 2018/05/16	53
Figure 19. Granulométrie de la zone d'étude, classifiée avec les images #6 et 7, du 2018/05/13	54
Figure 20. Granulométrie de la zone d'étude, classifiée avec l'image #5, du 2018/05/08	55
Figure 21. Granulométrie de la zone d'étude, classifiée avec l'image #4, du 2018/04/23	56
Figure 22. Granulométrie de la zone d'étude, classifiée avec les images #1,2 et 3, du 2016/04/21	57
Figure 23. Granulométrie de la zone d'étude, classifiée avec l'image #8, du 2018/05/16 (2 classes)	58
Figure 24. Granulométrie de la zone d'étude, classifiée avec les couvertures du 18/05/16, 18/05/21, 18/05/13 et 18/06/10	59
Figure 25. Granulométrie de la zone d'étude, calculée par agrégat (par mode) post-classification des couvertures du 18/05/16, 18/05/13 et 18/06/10	60
Figure 26. Photo de proximité d'une terre agricole, 16 mai 2019	61



## LISTE DES ABBRÉVIATIONS

### Longueurs d'onde :

B	Bleu
V	Vert
R	Rouge
PIR	Proche infrarouge
IRC	Infrarouge à courte longueur d'onde
LO	Longueur d'onde
LB	Longueur de bande
nm	Nanomètre

### Informatique :

GeoTIFF	<i>Geo Tagged Image File Format</i>
GEE	<i>Google Earth Engine</i>
ID	Identifiant

### Granulométrie :

A	Argile	LSG	Loam sableux grossier
ALi	Argile limoneuse	LSTF	Loam sableux très fin
Alo	Argile lourde	SL	Sable loameux
AS	Argile sableuse	STF	Sable très fin
L	Loam	STFL	Sable très fin loameux
LA	Loam argileux	SFL	Sable fin loameux
Li	Limon	S	Sable
LLi	Loam limoneux	SF	Sable fin
LLiA	Loam limono-argileux	SG	Sable Grossier
LS	Loam sableux	SGL	Sable grossier loameux
LSA	Loam sablo-argileux	Org	Organique
LSF	Loam sableux fin		

## **REMERCIEMENTS**

Je souhaite remercier Yacine Bouroubi, mon directeur de recherche, qui m'a soutenu et conseillé dans toutes les étapes du présent travail.

Je tiens aussi à remercier mes correcteurs, Mickaël Germain et Kalifa Goïta, qui ont mis en lumière des idées et des améliorations à faire que je n'avais pas envisagées.

## RÉSUMÉ

L'étude présente vise à tester l'utilisation d'images satellites Sentinel-2 multitudes pour la cartographie numérique de la texture de surface des sols agricoles. La zone d'étude est la Montérégie, principale région agricole du Québec, Canada. La recherche, le téléchargement et la correction atmosphérique des images (11 images Sentinel-2) ont été réalisés de manière semi-automatique à partir de *Google Earth Engine* et à l'aide de *GRASS GIS*. Des indices spectraux corrélés à la texture du sol, d'après la littérature scientifique, ont ensuite été calculés. L'entraînement de la classification des textures de sol a été effectué à l'aide d'une base de données d'échantillonnage des sols d'Agriculture et Agroalimentaire Canada, constituée de 37 517 points couvrant une région d'environ 5 000 km<sup>2</sup> de la zone d'étude. La méthode de classification utilisée est celle du maximum de vraisemblance. La différenciation des classes de granulométrie (sables, loams et argiles) n'a pas été très satisfaisante, avec un taux de succès d'environ 29,9% et un indice de Kappa de 0,32. Au niveau géographique, on obtient quelques informations intéressantes, notamment une certaine constance dans la localisation des classes. Le principal acquis pédagogique de ce projet est la réalisation d'une chaîne automatique de traitement d'images, basée sur des scripts *JavaScript* et *Python*. Cette chaîne de traitement comporte la recherche des images sur les serveurs de *Google Earth Engine*, leur téléchargement et leur prétraitement (corrections atmosphériques), qui permettent leur analyse (calcul d'indices spectraux et classification) et enfin la production d'une information utile (cartes de textures de sol).

## 1. INTRODUCTION

### 1.1. CONTEXTE

L'agriculture dans le monde s'est beaucoup transformée avec l'apparition de technologies de plus en plus avancées. La pratique se faisait auparavant sur des terres de tailles moins importantes, avec une plus grande proximité entre l'agriculteur et ses cultures ainsi qu'un rendement nettement inférieur à celui d'aujourd'hui. Grâce aux diverses découvertes technologiques et scientifiques que l'humanité a connues, les travaux de plantation, de fertilisation et de moisson se font sur de grandes surfaces, de façon mécanique et homogène, avec des rendements nettement supérieurs (Plant, 2001). L'essai présent tente d'apporter une information utile sur les sols agricoles dans ce contexte de changement et d'évolution.

La production agricole au Québec est une composante majeure de son économie et est constituée d'un peu moins de 30 000 exploitations. Sa part du marché est aux deux tiers occupée par l'élevage et au

tiers par la production végétale, qui est celle qui nous intéresse ici. En 2017, des 3 milliards de dollars brut générés par celle-ci (excluant l'acériculture), 3 cultures principales dominaient ; le maïs (21%), les légumes frais (19%) et le soya (16%), pour un total de 56% (MAPAQ, 2018).

Depuis les dernières décennies, l'agriculture dans le monde a beaucoup évolué. Il ne manque d'ailleurs pas de raisons pour justifier un travail d'amélioration de la production des terres agricoles, autant en augmentation de leur rendement qu'en réduction des ressources nécessaires à leur exploitation. Ceci est nécessaire pour des raisons économiques, mais aussi parce que l'expansion agricole et l'agriculture industrielle sont responsables, notamment, de nombreuses pertes d'habitats naturels, de pollution et de transformation radicale des paysages et des territoires (Capmourteres et al., 2018; Plant, 2001). En améliorant le rendement des sols et en diminuant les intrants, l'expansion est potentiellement réduite ou ralentie et les effets sur l'environnement sont allégés.

Avec l'utilisation de l'information spatiale de plus en plus accessible, la pratique plus « traditionnelle » s'est donc tournée vers une pratique d'agriculture de précision ou on prend davantage compte de l'hétérogénéité intra-parcellaire des terres agricoles. Il est possible de faire des prédictions, des modèles et des cartographies au sujet des propriétés des sols ainsi que des stress et de la santé des plantes. Cela est possible grâce à l'apport de la télédétection par satellites, avions et drones. Les produits dérivés de chacune de ces sources de données permettent des analyses intra-parcellaires avec des résolutions variables, allant du centimétrique, avec les drones, au métrique, avec les images prises par avion ou satellite (Huang et al., 2018; Mulla et Khosla, 2015).

Un point commun à ces trois types de plateformes est la façon dont les images sont exploitées ainsi que les bandes du spectre électromagnétique qui sont utilisées. On parle ici notamment des bandes bleue (environ 500 nm), verte (environ 550 nm), rouge (environ 660 nm) et proche infrarouge (environ 830 nm) (Mulla et Khosla, 2015).

L'agriculture de précision est un domaine qui, comme bien d'autres, se saisit bien par l'entremise d'exemples, comme le système d'aide à la décision en agriculture de précision *SCAN (Sol, Culture, Atmosphère pour N- (azote))*. Développé par Agriculture et Agroalimentaire Canada et offert commercialement par Effigis depuis 2017, ce système permet d'améliorer l'efficacité d'utilisation de l'azote (NUE) de deux façons; en réduisant l'utilisation de fertilisant azoté et en augmentant le rendement. Le système prend en compte notamment la texture du sol, les conditions météorologiques, les pratiques de gestion et les paramètres économiques (Effigis, 2019a). En chiffres, des économies annuelles proches de 50\$ par hectare sont possibles (Agriculture et Agroalimentaire

Canada, 2017). Avec des terres agricoles de plusieurs centaines d'hectares, ces montants deviennent substantiels.

Pour revenir à l'imagerie, des centaines de satellites sont présentement actifs en orbite autour de la Terre, une bonne partie d'entre eux servant à la télédétection (Huang et al., 2018). La télédétection par satellite, en contraste avec les avions ou les drones, est celle qui permet les études aux échelles géographiques les plus importantes par l'accessibilité des images et la couverture globale de celles-ci. Par exemples, les satellites Landsat-7 et -8, avec une résolution de 30 mètres, passent par-dessus l'entièreté de la planète avec une revisite de 16 jours, et avec une différence de 8 jours entre les deux, ce qui permet un repassage de 8 jours lorsqu'utilisés conjointement (USGS, s.d.). La constellation Sentinel-2, avec un capteur de résolution de 10 mètres, permet avec deux satellites (Sentinel-2 A et Sentinel-2 B) un repassage aux 5 jours (ESA, 2019b). Dans les deux cas, des bandes dans le visible, dans diverses longueurs d'onde de l'infrarouge et plus encore sont disponibles. Il s'agit ici de seulement 4 satellites, mais ils représentent à eux seuls une quantité phénoménale d'images. Cette disponibilité de données entre dans un nouveau paradigme de l'information, le *Big Data*, qui apporte beaucoup de possibilités mais aussi bien des défis, notamment la récupération et le traitement de données très nombreuses et volumineuses. Ces données peuvent être caractérisées avec les 3 V : *Volume*, *Variété* et *Vélocité*, soit respectivement la quantité de données, leur nature de plus en plus hétérogène ainsi que la rapidité de renouvellement et de production de celles-ci (Lokers et al., 2016).

Cette étude a donc sa place dans un contexte d'agriculture de précision en évolution, d'abondances de données de télédétection et de comment ces domaines peuvent être liés efficacement.

## 1.2. PROBLÉMATIQUE

L'agriculture de précision demandant que l'on connaisse bien les caractéristiques des sols, on peut se poser la question suivante : cette tâche est-elle remplie, au Québec? Des données de pédologie l'IRDA sont disponibles, celles-ci étant cependant à l'échelle de 1:20 000 (IRDA, 2019), elles ne permettent pas à elles seules la gestion intra-parcellaire des terres. Une donnée ouverte représentant la pédologie à une échelle détaillée sur une couverture constante n'existe pas et il serait pertinent de développer une méthode pour en produire grâce à la télédétection, qui évite ou réduit potentiellement le travail de terrain. La granulométrie étant une caractéristique importante du sol lorsqu'il est question de sa gestion et/ou de son exploitation (Agriculture et Agroalimentaire Canada, 2014), il s'agit de l'aspect que l'on va tenter d'évaluer.

### 1.3. OBJECTIFS

Cela nécessite d'obtenir des images facilement traitables (couverture nuageuse faible/absente) et dont les sols agricoles sont nus. De plus, puisqu'il s'agit de caractériser les sols à l'aide de leurs réflectances, il est judicieux d'effectuer des corrections atmosphériques sur les images obtenues. Dans un contexte actuel de données sont volumineuses et nombreuses, il serait nécessaire d'effectuer ces deux tâches de façon automatique ou semi-automatique.

L'objectif de ce travail donc est la cartographie des caractéristiques granulométrique des sols agricoles à partir d'imagerie optique Sentinel-2 en Montérégie. Pour en arriver à cette fin, nous avons défini 4 objectifs secondaires :

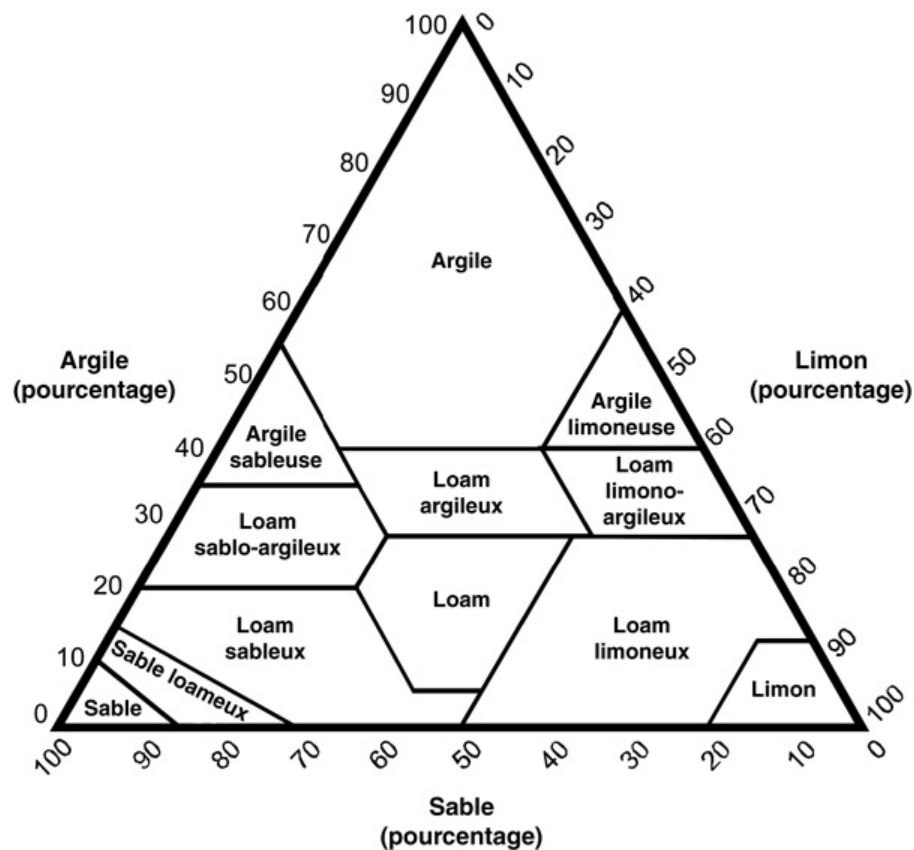
- développer une méthode permettant de récupérer des images Sentinel-2 selon leur localisation, date d'acquisition et couverture nuageuse;
- développer une méthode permettant d'effectuer les corrections atmosphériques sur les images obtenues;
- développer une classification en croisant les données optiques aux données *in situ* de granulométrie des sols;
- effectuer une analyse qualitative et statistique des données matricielles produites ainsi qu'une cartographie thématique associée.

Ultimement, l'intérêt est d'étudier la possibilité d'utiliser les indices spectraux spécifiques à la caractéristique d'un sol afin d'automatiquement lui assigner une texture.

## 2. CADRE THÉORIQUE

### 2.1. GRANULOMÉTRIE DES SOLS

La granulométrie est un aspect très important du sol lorsqu'il est question d'agriculture (Bouroubi et al., 2016; Agriculture et Agroalimentaire Canada, 2014). Les sols argileux et limoneux sont connus pour être plus faciles à manipuler tandis que les sols sableux sont plus légers et malléables. Les sols argileux absorbent moins rapidement l'eau et sont plus propices au ruissellement. Les sols limoneux sont sensibles à l'érosion par l'eau. Les sols sableux absorbent l'eau plus vite, ce qui peut être un problème lors de l'utilisation intensive de pesticides, ceux-ci étant rapidement propagés dans le sous-sol, ils sont aussi plus propices à l'érosion éolienne. Il est important pour un agriculteur de connaître la texture de sa ou ses terre(s) et d'agir en conséquence (Agriculture et Agroalimentaire Canada, 2014). La granulométrie est illustrée par le triangle des sol (Figure 1) qui donne les pourcentages d'argile, limon et sable contenus dans le sol.



Source : OMAFRA, 2012

Figure 1. Triangle des textures de sol (version canadienne)

## 2.2. TÉLÉDÉTECTION ET CARTOGRAPHIE DES SOLS

Il peut être légitime de se poser au départ la question à savoir si l'utilisation de la réflectance peut être utile pour détecter des caractéristiques granulométriques/texturales. Non seulement elle peut l'être; il y a une relation notable entre la granulométrie et la réflectance dans l'infrarouge et le visible (Clément et al., 1983). Il n'est d'ailleurs pas toujours certain que des analyses en laboratoire soient davantage adaptées; en plus d'être plus coûteuses et lentes, elles offrent dans certains cas des résultats inférieurs que ceux obtenus par évaluation de la réflectance dans les bandes visibles, proche infrarouge et infrarouge à courte longueur d'onde (Gholizadeh et al., 2016). Dans le même ordre d'idée, une étude de Hummer et al. (2001) démontre l'efficacité de l'estimation de l'humidité et de quantité de matière organique dans le sol lors d'analyses de proximité usant du proche infrarouge. Certainement il n'est pas question ici d'affirmer que l'évaluation de la réflectance est supérieure que l'analyse en laboratoire, mais qu'elle peut déjà constituer un substitut dépendamment des situations.

Il s'agit ci-haut de la détection de proximité, qu'en est-il de la télédétection? Des indices spectraux ont fait leur preuve pour ce qui est de détecter des caractéristiques minéralogiques et texturales des sols; on parle notamment du *Redness Index*, du *Texture Index* (ou *Normalized Difference Moisture Index*) et du *Color Index* (Houssa, et al., 1996; Bouroubi et al., 2016). Sont aussi pertinents le *Saturation Index* et le *Brightness Index* (Bouroubi et al., 2016) ainsi que l'*Organic Matter Index* (Gili et al., 2017), bien que dans une mesure différente pour le dernier. Dans le cas de la télédétection, ces indices sont calculés par valeurs de réflectance (valeurs de pixels), comme au Tableau 1.

Tableau 1. Indices spectraux pertinents

Indice	Formule
<i>Redness Index</i>	$R^2 / (B \times V^2)$
<i>Coloration Index</i>	$(R - V) / (R + V)$
<i>Organic Matter Index</i>	$1 / V^2$
<i>Normalized Difference Moisture Index</i>	$(PIR - IRC) / (PIR + IRC)$
<i>Brightness Index</i>	$\sqrt{[(B^2 + V^2 + R^2) / 3]}$
<i>Saturation Index</i>	$(R - B) / (R + B)$

Où : V = vert, B = bleu, R = rouge, PIR = proche infrarouge et IRC = infrarouge à courte longueur d'onde

Dans certains cas, même si une faible végétation est présente sur le sol, les parcelles agricoles peuvent être vues comme étant en sol nu sur l'image satellite. Cette faible végétation influencera la réflectance



du sol et par conséquent les indices spectraux dérivés de l'image. La végétation morte (résidus de cultures) pose un problème sensiblement moindre, mais tout de même non négligeable (Siegal et Goetz, 1977).

### 2.3. SERVICES WEB D'IMAGERIE

Les images satellites étant de plus en plus nombreuses et de plus en plus disponibles, explorer les possibilités de les télécharger automatiquement devient plus complexe. Il existe plusieurs services web qui permettent de faire cela de façon plus ou moins conviviale, notamment :

- la plateforme *Earth*, intégrée à *Amazon Web Service* (Amazon, 2019), qui propose un service payant qui permet de trouver, traiter et télécharger des images satellitaires (notamment Landsat-8 et Sentinel-2). Le service est accessible à l'aide du langage *Python* mais non sur navigateur;
- la plateforme *Radian Earth*, qui propose un service gratuit ou payant sur navigateur qui permet de sélectionner et télécharger des images de plusieurs satellites, comme Sentinel-2, Landsat-4, -5, -7 et -8, Worldview, etc. (Radian Earth Foundation, 2019);
- *Google Earth Engine* est un service de Google (2019b) sur navigateur qui propose de sélectionner, traiter et télécharger des images à l'aide de scripts en *JavaScript*. Comme les autres, il possède des données de plusieurs missions satellitaires. Il est spécifiquement intéressant parce qu'il est entièrement gratuit, tant que le travail qu'on y effectue est à des fins éducatives ou de recherche scientifique. Dans le cadre de cette étude, il est celui qui nécessite le moins d'installation logicielle, c'est-à-dire aucune, si ce n'est que l'inscription au service et l'utilisation d'un script pour accéder aux données.

### 2.4. CORRECTION ATMOSPHERIQUE

Afin de produire un modèle fiable, il est important que les images représentent au maximum la réalité au sol. Ce n'est pas le cas avec les images satellites brutes; la valeur numérique que les capteurs enregistrent n'est pas parfaite (voir Figure 2). Celle-ci est affectée de plusieurs façons (Ose et al., 2016; Kergomard, 1996), notamment par l'atmosphère. Ainsi :

- les rayons du soleil sont faiblement reflétés par l'atmosphère en y pénétrant;
- ils sont ensuite absorbés et diffusés partiellement en la traversant;
- en frappant le sol, ils sont partiellement transmis, absorbés et reflétés de nouveau;

- les rayons qui au sol repassent par l'atmosphère, ils y sont donc absorbés et diffusés une seconde fois;
- ils atteignent finalement le capteur, proposant une image proche de la réalité mais non entièrement conforme.

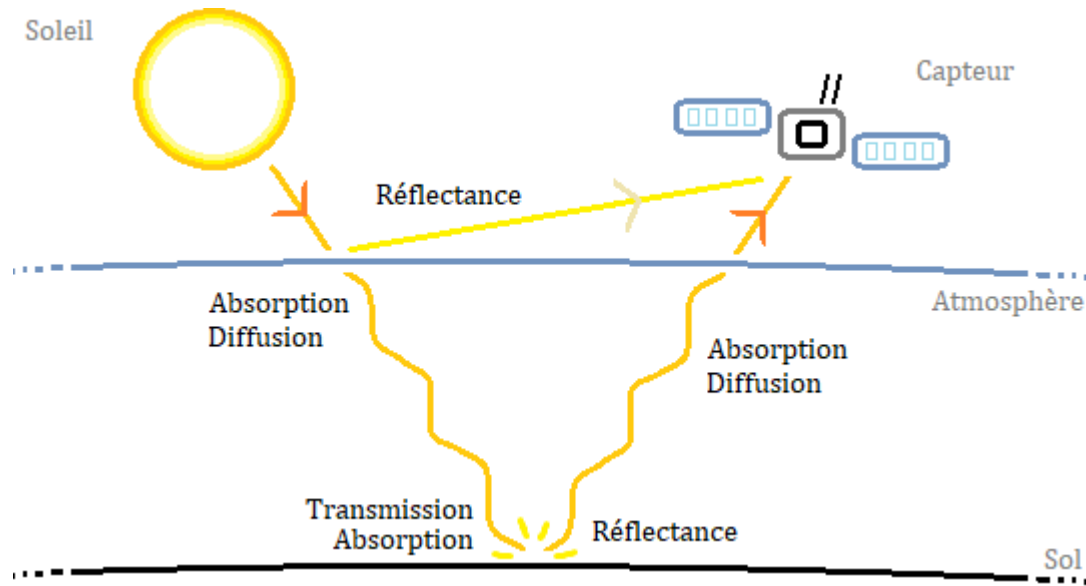


Figure 2. Trajet de la lumière captée par satellite

La correction atmosphérique est donc un prétraitement important à accomplir si l'analyse des données nécessite une grande précision (Sola et al., 2018; Bouroubi et al., 2018).

Considérant que le contenu et la densité de l'atmosphère sont variables dépendamment de la localisation, il est extrêmement fastidieux de créer un modèle de correction entièrement adapté à la situation, mais il est possible d'utiliser des modèles représentant des situations générales. Des résultats de correction approximatifs sont donc relativement facilement accessibles (Ose et al., 2016).

Plusieurs modèles/algorithmes de corrections existent pour les données satellitaires. L'étude de Sola et al. (2018) en compare quatre pour ce qui est des images Sentinel-2 : (1) *SEN2COR*, un processus de correction qui utilise une détection des nuages et des vapeurs d'eau ainsi qu'une classification de la scène, (2) *MAJA*, un outil qui sert à corriger des séries temporelles avec des prises de vue équivalentes, (3) *iCOR*, un processus fondé sur l'identification des pixels au sol et sur l'eau et (4) *6S* (Seconde Simulation du Signal Satellitaire dans le Spectre Solaire), un outil libre d'utilisation qui utilise des modèles atmosphériques et d'aérosols. Toutes ces méthodes tiennent compte de l'angle d'acquisition

et de l'angle solaire au moment de la prise de vue. Parmi celles-ci, les plus efficaces sont, selon les résultats de cette étude, *MAJA* et *iCOR* pour ce qui est de l'évaluation de la réflectance au sol. Les différences sont cependant minces entre toutes les méthodes et elles produisent toutes un résultat supérieur aux images d'origine dans toutes les bandes spectrales.

Le modèle 6S ne brille pas particulièrement dans l'étude de Sola et al. (2018), il s'agit cependant d'un modèle établi qui peut produire des corrections fiables avec peu de données en entrée (Sharma et al., 2009; Vermote et al., 1997; Zhao et al., 2001). Le modèle est une continuation du modèle 5S (Simulation du Signal Satellitaire dans le Spectre Solaire), développé en 1992 puis retravaillé à de nombreuses reprises (Vermote et al., 2006; Bouroubi et al., 2018).

Il s'agit notamment du modèle utilisé par le module *i.atcorr* de *GRASS Gis*, un logiciel libre d'accès. Cela n'est pas idéal et nécessite certaines approximations, mais le module peut être au minimum utilisé avec comme seule donnée d'entrée l'image à corriger en question (OSGeo4W, 2019). C'est ce modèle qui sera utilisé dans notre travail, pour des raisons de fiabilité reconnue mais aussi d'accessibilité.

### 3. MATÉRIEL ET MÉTHODES

La Figure 3 suivante montre brièvement les principales étapes du présent travail :

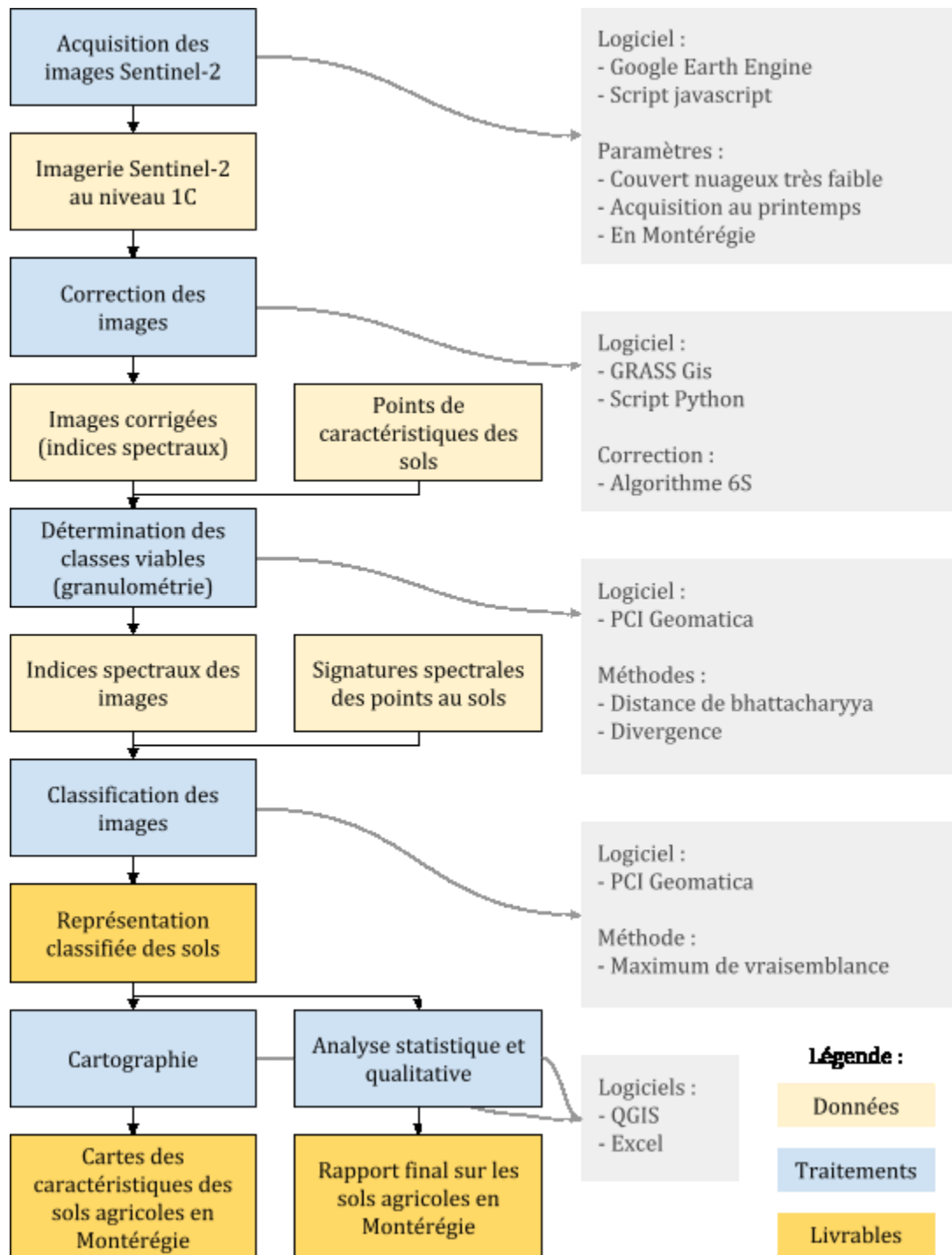


Figure 3. Étapes principales du présent travail

### 3.1. SITE D'ÉTUDE

Le site d'étude sélectionné (Figure 4) est situé en Montérégie, surnommée le « grenier du Québec ». Cette région administrative, située à l'est de Montréal et à l'ouest du Centre-du-Québec et de l'Estrie, fait un peu moins de 12 000 km<sup>2</sup>.

Son territoire est principalement d'occupation agricole intensive (grandes cultures et exploitations laitières), agroforestière et forestière (Jobin et al., 2004). Établir la zone dans cette région est idéal puisque 86% du territoire est couvert d'exploitations agricoles, dont 60% sont utilisées pour la production végétale. Cela constitue le quart des terres cultivées au Québec (MAPAQ, 2014). La zone constitue une aire qui peut être entièrement couverte par seulement quelques images satellitaires, dépendamment de la fauchée de celles-ci. Dans le cas des images Sentinel-2, comme il le sera détaillé plus bas, 3 images sont suffisantes pour entièrement couvrir la zone.

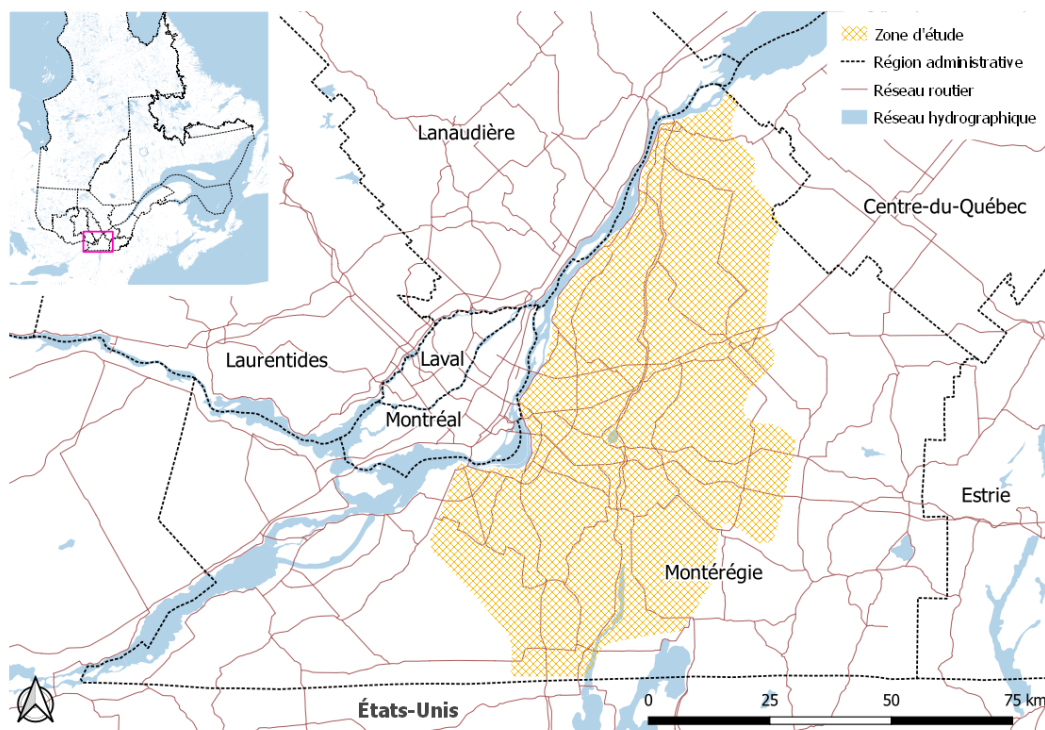


Figure 4. Zone d'étude

Source des données routières, hydriques et administratives : MERN, 2013

La superficie de la zone d'étude fait un peu moins de 5 000 km<sup>2</sup>. Elle est définie selon deux critères : (1) la localisation de l'étude, qui vise la région de la Montérégie; et (2) la disposition géographique des données d'échantillonnage de sol utilisées pour classer les images (voir 3.2.2).

## 3.2. DONNÉES

Cette étude nécessite des données de deux natures : de l'imagerie satellitaire au format matriciel, afin d'effectuer la classification des sols, ainsi que des points d'échantillonnage *in situ* des caractéristiques des sols au format vectoriel, afin d'entraîner ladite classification. Ces données permettent ultimement, une fois les divers corrections et traitements effectués, d'obtenir des produits avec lesquels des travaux d'analyse et de cartographie sont possibles. Mis à part la correction atmosphérique et la classification finale des images, les données sont prétraitées et traitées dans QGIS (OSGeo4W, 2019b).

Ultimement, il ne s'agira pas ici de *Big Data*, considérant la taille modeste de la zone d'étude et de la temporalité, mais l'esprit du travail est principalement dans le concept de l'automatisation, et non dans la taille des données utilisées finalement.

### 3.2.1. Imagerie

Les images utilisées proviennent de la mission Sentinel-2, de l'Agence spatiale européenne. Le capteur utilisé, *MSI*, pour *MultiSpectral Instrument*, prend des images sur une fauchée de 290 km avec un période d'observation de 17 minutes en moyenne (ESA, 2019b).

Ces données sont du niveau de prétraitement 1C, ici acquises à l'aide du service *Google Earth Engine* (Google, 2019b). Une scène de niveau 1C a déjà passé par plusieurs processus de corrections radiométriques et géométriques mais pas par la correction atmosphérique (ESA, 2019a). L'étape d'acquisition de ces images est une partie primordiale de la présente méthode et est décrite en 3.3.1.

Des informations détaillées sur les bandes du spectre électromagnétique du capteur *MSI* sont disponibles au Tableau 18, en annexe. Dans le cadre de ce travail cependant, seules les bandes 2,3,4,8 et 12 (Bleu, vert, rouge, proche infrarouge et infrarouge à courte longueur d'onde) sont utilisées dans les traitements. Les autres bandes sont ignorées, ce qui est un peu opposé au concept de *Big Data* qui s'applique idéalement dans un contexte où les ressources informatiques (vitesses de traitements et capacités en stockage) sont plus importantes. Cette simplification allège dans le cas présent énormément la taille des données nécessaires et accélère de façon notable les nombreux traitements à accomplir. Le Tableau 2 présente la liste des images utilisées, selon leur identifiant sur *Google Earth Engine*.

Tableau 2. Données Sentinel-2 utilisées

1	ID_GEE	COPERNICUS/S2/20160421T160034_20160421T194244_T18TWQ
	Date	21/04/2016 16:00:34
2	ID_GEE	COPERNICUS/S2/20160421T160034_20160421T194244_T18TWR
	Date	21/04/2016 16:00:34
3	ID_GEE	COPERNICUS/S2/20160421T160034_20160421T194244_T18TXQ
	Date	21/04/2016 16:00:34
4	ID_GEE	COPERNICUS/S2/20180423T153909_20180423T154331_T18TXR
	Date	23/04/2018 15:43:31
5	ID_GEE	COPERNICUS/S2/20180508T153911_20180508T154947_T18TXR
	Date	08/05/2018 15:49:47
6	ID_GEE	COPERNICUS/S2/20180513T153809_20180513T154810_T18TXR
	Date	13/05/2018 15:48:10
7	ID_GEE	COPERNICUS/S2/20180513T153809_20180513T154810_T18TXS
	Date	13/05/2018 15:48:10
8	ID_GEE	COPERNICUS/S2/20180516T154809_20180516T155840_T18TXR
	Date	16/05/2018 15:58:40
9	ID_GEE	COPERNICUS/S2/20180521T154911_20180521T155356_T18TXR
	Date	21/05/2018 15:53:56
10	ID_GEE	COPERNICUS/S2/20180531T155021_20180531T155122_T18TXR
	Date	31/05/2018 15:51:22
11	ID_GEE	COPERNICUS/S2/20180610T154901_20180610T155519_T18TXR
	Date	10/06/2018 15:55:19

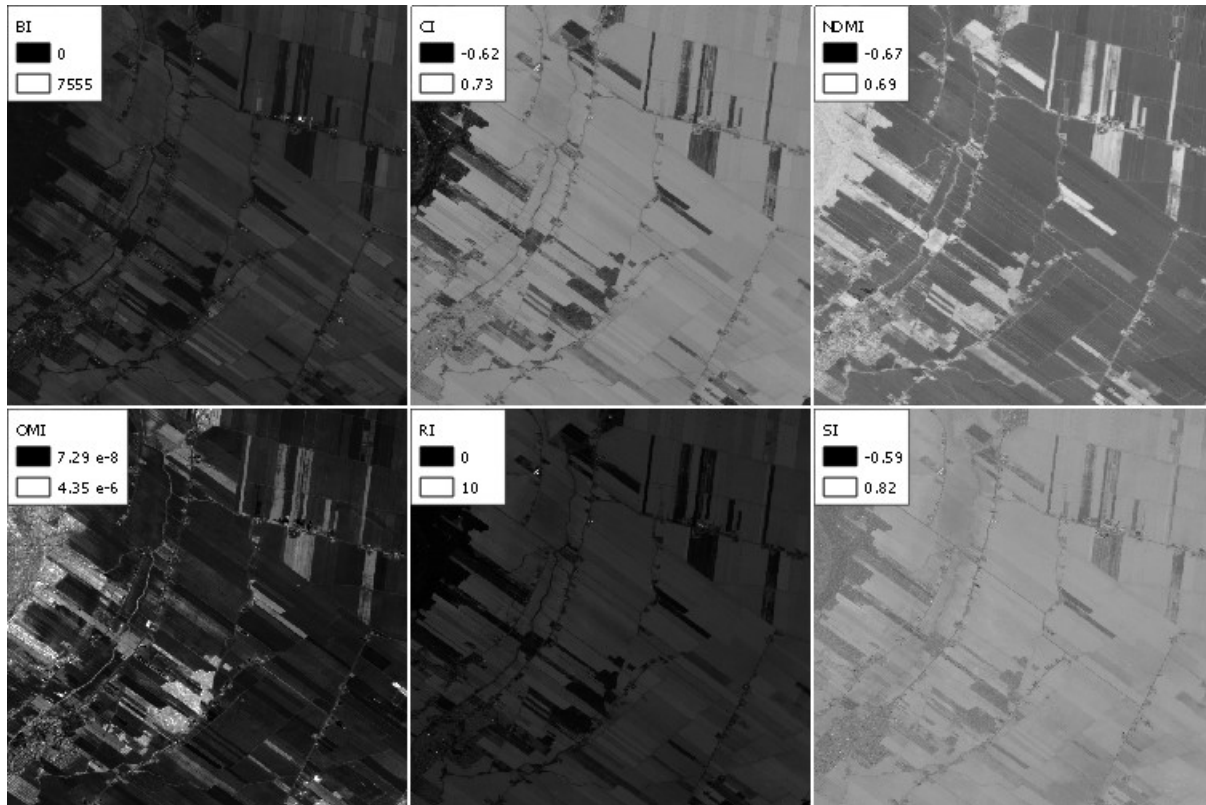
Où : Id GEE = identifiant sur Google Earth Engine, Date = Moment d'acquisition en temps UTC

Les granules 1, 2 et 3 du tableau précédent sont géographiquement côtes à côtes et possèdent les mêmes temps d'acquisition et identifiant d'origine, il s'agit donc d'une image séparée par le service de Google. Une fois téléchargées ces images sont alors fusionnées pour les traitements qui suivent. Il en est de même pour les granules 6 et 7. Les identifiants originaux des granules sont disponibles en annexe, au Tableau 17.

Toutes ces images ont une couverture nuageuse si faible qu'elle n'est pas considérée, sauf la #10, celle-ci ayant une couverture légèrement problématique. Étant suffisamment localisée, celle-ci a été coupée à l'aide de polygones vectorisés à la main. La couverture géographique des scènes est variable; certaines couvrent presque l'entièreté de la zone, d'autres qu'une petite section. Les points de référence sont donc coupés en fonction des images (voir la Figure 7 à la section 3.2.2 suivante).



Ces données matricielles ont pour but ultime d'être utilisées pour créer les indices spectraux montrés au Tableau 1, soit : *Redness Index*; *Normalized Difference Moisture Index*; *Coloration Index*; *Brightness Index*; *Organic Index*; *Saturation Index* (la Figure 5 suivante les présente rapidement). Ce sont ces indices spectraux qui sont utilisés pour la classification des textures de sol.



À partir de l'image du 21/05/2018. Où : BI = *Brightness Index*, CI = *Coloration Index*, NDMI = *Normalized Difference Moisture Index*, OMI = *Organic Matter Index*, RI = *Redness Index*, SI = *Saturation Index*

Figure 5. Représentation visuelle des indices spectraux avec zoom sur une zone aléatoire

### 3.2.2. Points d'échantillonnage de sol

Les données des caractéristiques proviennent d'Agriculture et Agroalimentaire Canada (2013) et sont au format *shapefile* (points). Ces points sont au nombre de 60 944 au départ. L'information qui nous intéresse est la granulométrie (texture de surface). Celle-ci est séparée en 23 classes, certaines étant très peu représentées, d'autres en grand nombre. Le Tableau 3 les présente.



Tableau 3. Points de granulométrie *in situ*

Code	Description	Nombre	%
A	Argile	2 564	4,21
ALi	Argile limoneuse	9526	15,63
Alo	Argile lourde	406	0,67
AS	Argile sableuse	18	0,03
L	Loam	13 574	22,27
LA	Loam argileux	5 380	8,83
Li	Limon	9	0,01
LLi	Loam limoneux	2 294	3,76
LLiA	Loam limono-argileux	7 948	13,04
LS	Loam sableux	4 769	7,83
LSA	Loam sablo-argileux	492	0,81
LSF	Loam sableux fin	3 789	6,22
LSG	Loam sableux grossier	60	0,10
LSTF	Loam sableux très fin	569	0,93
SL	Sable loameux	2 229	3,66
STF	Sable très fin	29	0,05
STFL	Sable très fin loameux	106	0,17
SFL	Sable fin loameux	2 408	3,95
S	Sable	449	0,74
SF	Sable fin	937	1,54
SG	Sable Grossier	16	0,03
SGL	Sable grossier loameux	30	0,05
Org	Organique	3 342	5,48

Les classes en jaune ci-haut sont celles dont le nombre s'est révélé insuffisant pour les utiliser. Il reste néanmoins, après suppression, 57 823 points, soit 94,88 %. Les 3 121 points perdus représentent 5,12 % de l'ensemble. Les sols organiques sont suffisamment représentés, mais pour des raisons de différenciation de classe (voir section 3.3.3), ils ne sont pas utilisés pour la classification finale.

Certains de ces points ne sont pas toujours situés à des endroits correspondants aux sols nus ou agricoles (bien que la majorité le soit, comme à la Figure 6). Les points problématiques sont surtout situés soit en zones végétalisées ou urbaines, cela est une contrainte pour l'établissement des paramètres de classification puisque la nature de ces sols et leur réflectance diffèrent de ce qui est recherché.

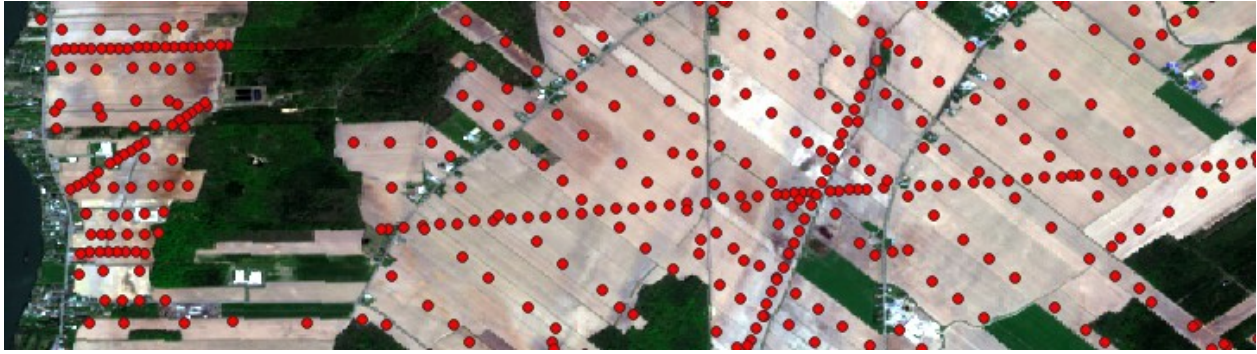


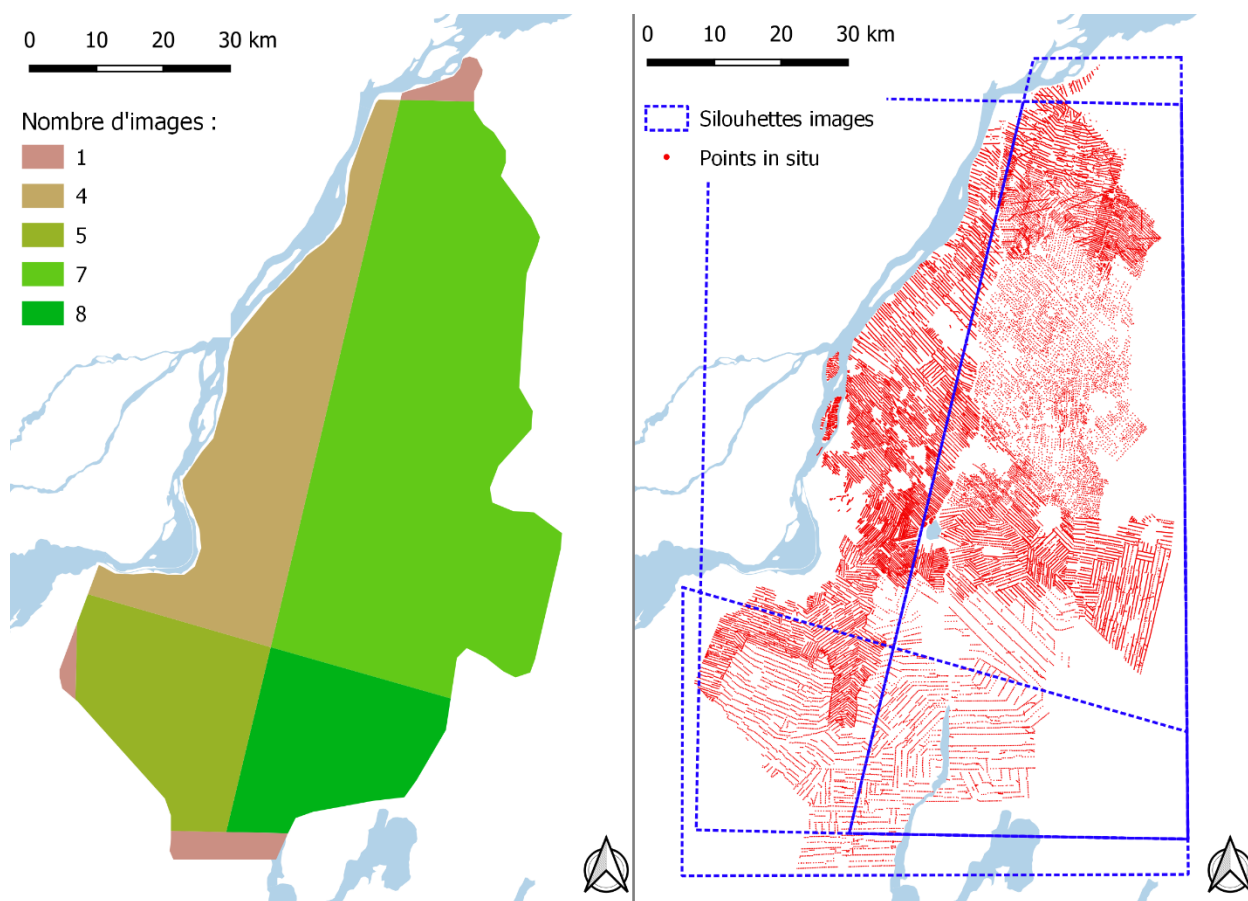
Figure 6. Points d'échantillonnage *in situ* du sol (agrandissement aléatoire sur la zone d'étude)

Pour éviter ce problème, il suffit de localiser et de ne pas utiliser ces points. Cela s'est fait de deux façons :

- les points situés en zones végétalisées sont exclus en prenant compte du NDVI calculé avec l'imagerie. Seuls les points correspondants à un NDVI inférieur à 0,35 sont compris. Cette valeur a été sélectionnée puisque légèrement au-dessus du NDVI maximum que les terres agricoles produisent dans ce cas-ci. Il s'agit ici d'un compromis, car bien que de la végétation (gazon, petites plantes, pousses, etc.) peut être toujours présente pour un NDVI de 0,35, comme à la Figure 26 en annexe, ce seuillage est choisi afin de conserver un nombre satisfaisant de points de comparaison. La considération d'un seuil de NDVI de 0,2 élimine pratiquement tous les points;
- les points situés en territoires urbains sont exclus à l'aide de polygones représentant ces zones, vectorisés à la main.

Une fois les données non désirées supprimées, on se retrouve avec 37 519 points, soit 61,56 % des points originaux, ce qui constitue une perte substantielle, mais non fatale.

Comme mentionné plus haut, chaque image couvre partiellement la zone d'étude, ce qui ne permet pas d'effectuer la comparaison images/points de la même façon pour chacune d'entre elles. De plus, les points ne sont pas géographiquement disposés de façon homogène. Comme on peut cependant le voir sur la Figure 7, cela n'empêche pas que chaque couverture comporte un nombre important de points.



À gauche, nombre d'image(s) par couverture selon l'étendue des points. À droite, distribution géographique des points.

Figure 7. Zone d'étude selon la distribution des points et le nombre d'images

Le Tableau 4 suivant présente le nombre de points par couverture (date) après nettoyage complet, c'est-à-dire que les classes de granulométrie trop peu représentées ont été enlevées et que seuls les points en sols nus sont compris.

Tableau 4. Nombre final de points par couverture

Date	21/04/2016			23/04/2018	08/05/2018	13/05/2018		16/05/2018	21/05/2018	31/05/2018	10/06/2018
Num.	1	2	3	4	5	6	7	8	9	10	11
Points	6 365			22 790	21 343	20 503		34 078	34 078	25 786	35 601

Des zones tampons de 10 x 10 mètres (même résolution spatiale que les données satellites) ont été créées autour des points afin qu'ils puissent être utilisés comme sites d'entraînement dans *Geomatica* (PCI Geomatics, 2019) pour la classification.

### 3.3. MÉTHODOLOGIE

#### 3.3.1. Acquisition semi-automatique des images Sentinel-2

Les images Sentinel-2 sont obtenues au travers du service *Google Earth Engine* (voir Google (2019b)). La requête d'imagerie se fait en *JavaScript*, sur le *Code Editor* du service, directement sur navigateur. Elle interroge les métadonnées des images afin de satisfaire 3 paramètres et n'obtenir que des données pertinentes :

- une couverture nuageuse inférieure à 10%, afin de réduire l'occlusion. Cette valeur est récupérée à partir de l'évaluation de ladite couverture présente dans les métadonnées des images;
- une fenêtre temporelle limitée, soit entre le 15 avril et le 15 juin, peu importe l'année. Cela est nécessaire puisque pour observer les caractéristiques de sols agricoles, ceux-ci doivent être nus, donc sans végétation et non couverts de neige/glace (voir les Figures 14 et 26, en annexe);
- une couverture géographique correspondante à la zone d'étude;
- la récupération des bandes rouge, verte, bleue, proche infrarouge et infrarouge à courte longueur d'onde, cela n'est pas obligatoire ni même recommandé de façon générale mais permet d'alléger les téléchargements et traitements.

Les images trouvées sont affichées et peuvent être examinées avant d'être téléchargées (Figure 8) :

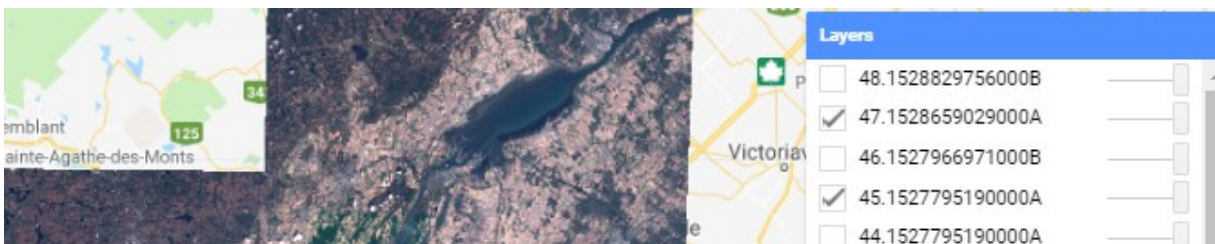


Figure 8. Affichage des images après exécution du script de sélection sur *Google Earth Engine* (Capture d'écran partielle)

Les images choisies sont ensuite découpées afin de ne représenter que la zone voulue, exportées au format *GeoTIFF* dans le compte *Drive* correspondant (voir Google (2019a)) avec comme nom de fichier le temps d'acquisition. Elles peuvent ensuite être téléchargées sur disque local afin de pouvoir être utilisées librement.

Ce script a été conçu pour être utilisé en complément avec le script suivant, soit celui de correction atmosphérique. Il permet de rapidement sélectionner les images pertinentes à un besoin particulier, il est disponible en annexe, au « Document 1 ».

### 3.3.2. Correction atmosphérique semi-automatique des images

La correction se fait à l'aide du module *i.atcorr*, intégré dans le logiciel *GRASS Gis* (OSGeo4W, 2019). Le module utilise l'algorithme de correction atmosphérique 6S (Seconde Simulation du Signal Satellitaire dans le Spectre Solaire). Le module peut être utilisé de façon entièrement manuelle, mais cela est inefficace dans un contexte de mégadonnées ou même seulement un avec un nombre relativement important de données.

Dans le cas présent, le nombre d'images (11), avec 5 bandes, rend le travail très demandant en temps, fastidieux et susceptible aux erreurs d'inattention. Un script Python a donc été développé pour effectuer la correction de façon semi-automatique, limitant le risque d'erreur de manipulation.

Le script est conçu pour corriger les images Sentinel-2 spécifiquement, bien qu'il puisse être modifié pour corriger de l'imagerie provenant d'autres capteurs. Il se comporte en 4 étapes principales :

1. D'abord, l'utilisateur doit indiquer au début du script certaines informations. Il s'agit ici, en théorie, de la seule étape du processus où un contrôle est nécessaire :
  - le dossier de provenance des images et le dossier de sortie désiré;
  - le format des images;
  - le modèle atmosphérique parmi une liste d'options;
  - le modèle d'aérosols parmi une liste d'options;
  - le nombre de bandes, 5 (R, V, B, PIR et IR à courte longueur d'ondes) ou 13 (toutes les bandes);
  - la visibilité en kilomètres;
  - l'élévation moyenne.
2. La section suivante du script se charge de trouver les images selon le format mentionné et les importe dans le projet;
3. Une fois les images importées, le logiciel doit établir les paramètres 6S correspondants aux images. C'est là que le code précédent d'acquisition a une relative importance (en donnant automatiquement aux fichiers comme nom leur temp d'acquisition). Cela est primordial puisque cette section du présent script utilise ce nom pour remplir certains

paramètres, soit le temps d'acquisition ainsi que le satellite (Sentinel-2 A ou Sentinel-2 B). Les autres paramètres proviennent des informations que l'utilisateur a entré précédemment et de la géographie de l'image elle-même. Les paramètres sont automatiquement rentrés dans un fichier texte temporaire que *GRASS Gis* lit et utilise pour corriger chaque bande séparément.

4. Les bandes corrigées sont ensuite regroupées puis exportées dans le dossier de sortie désigné. Elles sont alors disponibles et peuvent être utilisées dans un autre logiciel SIG, si désiré.

Chaque bande est corrigée séparément dans le logiciel puisqu'un composant des paramètres 6S est le numéro de celle-ci, ce qui nécessite de créer un fichier texte pour chacune d'elles. Par exemple, la bande proche infrarouge d'une image utilisée dans ce projet est corrigée suivant la liste de consignes présentées à la Figure 9 suivante :

1	25
2	4 21 16.00 -73.6851 45.1722
3	2
4	3
5	90
6	35
7	-1000
8	173

Figure 9. Fichier .txt de paramètres 6S pour une bande d'une image

La structure du fichier va comme suit :

- la ligne 1 indique au logiciel le satellite utilisé. Dans ce cas, Sentinel-2 A. Cela est important puisque les capteurs sont différents d'un satellite à l'autre;
- la ligne 2 est utilisée pour déduire les angles solaires et de prise de vue. Dans ce cas, l'image a été prise précisément à 16h00 le 21 avril, en Montérégie évidemment, mais plus précisément aux alentours de la coordonnée -73.6851 (longitude) et (45.1722 latitude);
- la ligne 3 est le numéro du modèle atmosphérique prédéfini (celui-ci équivaut à « été en mi-latitude »);
- la ligne 4 est le numéro du modèle d'aérosols prédéfini (celui-ci équivaut à « urbain »);



- la ligne 5 est la visibilité, en kilomètres. Dans ce cas, celle-ci est une approximation et n'est pas fondée sur des données disponibles, cette information étant très difficile à obtenir. Cela bien entendu joue sur la précision de la correction des images;
- la ligne 6 est la hauteur moyenne, en mètres. Dans le cas présent, la valeur de 35 a été établie séparément selon la moyenne de l'altitude de la zone d'étude avec le MTN de *Google Earth Engine*;
- la ligne 7 est la hauteur du capteur, la valeur par défaut lorsque celui-ci est monté sur satellite, comme ici, est de -1000, indiquant que le capteur se trouve à l'extérieur de l'atmosphère;
- la ligne 8 est le numéro de la bande à corriger, chaque bande de chaque capteur ayant un numéro particulier. La bande proche infrarouge prise par le capteur du Sentinel-2 A est indiquée par le numéro 173.

Tous ces paramètres peuvent être changés dépendamment des besoins. La Figure 10 suivante montre un exemple visuel de la correction effectuée :

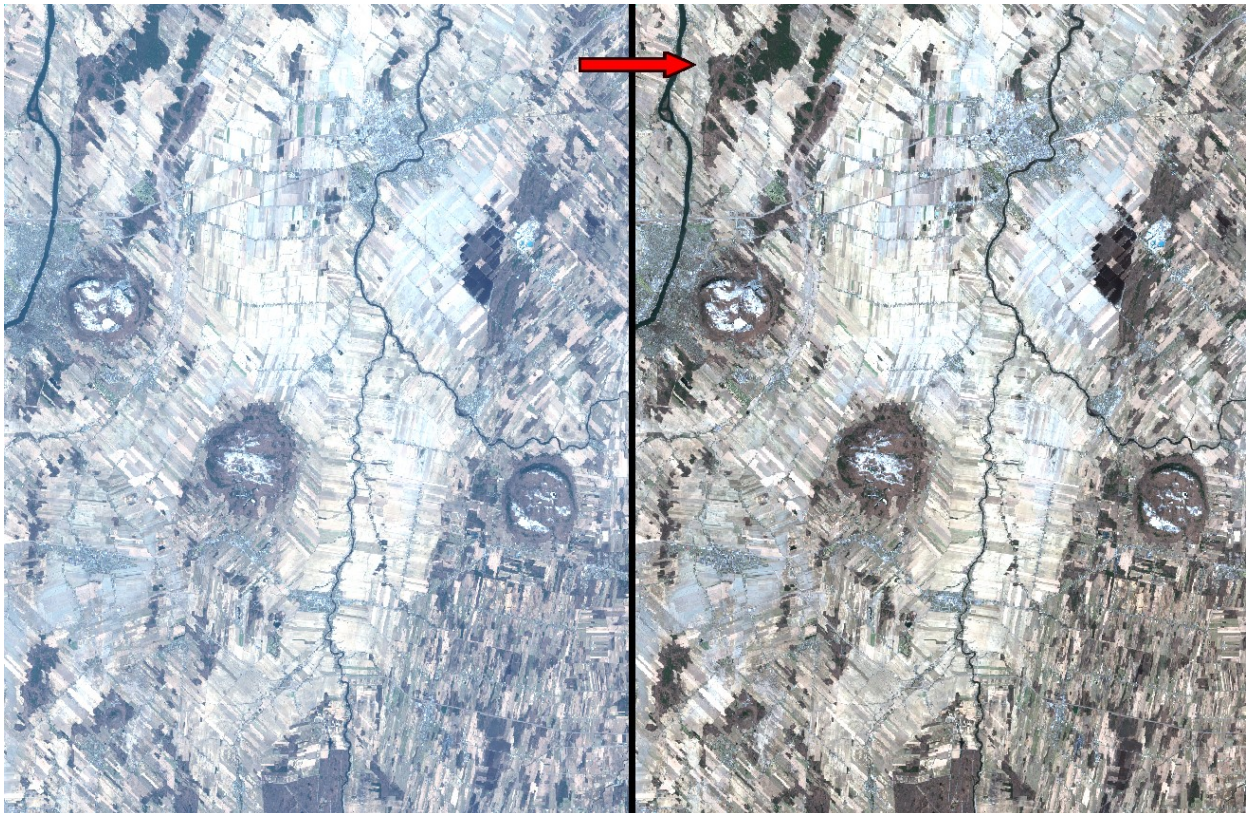


Figure 10. Agrandissement sur l'image #4 (2018/04/23), avant et après correction atmosphérique

L'entièreté du code Python est disponible au « Document 2 », en annexe.

### 3.3.3. Classification

Comme mentionné lors de la description des données *in situ*, celles-ci sont partiellement constituées de points dont la classe de sol est sous-représentée. Ces classes ont donc été enlevées puisque le nombre élevé de classes rendait la classification très brouillée, les différents types de sols étant mal différenciés par le logiciel en raison de leur faibles séparabilité spectrale (Tableau 5 suivant). La différenciation des classes a été établie à partir de la distance de Bhattacharyya et la divergence transformée, qui sont des indicateurs de séparabilité de classe. Ils sont compris entre 0 et 2 dans le cas présent, des valeurs de plus de 1,9 indiquent d'excellentes séparabilités (PCI Geomatics, 2018).

Tableau 5. Distance de Bhattacharyya avec l'entière des classes (sauf organique) avec l'image #9 de l'étude

	S	SF	SGL	SL	SFL	STFL	SFL	LS	LSF	Li	L	LLi	LSA	LA	LLiA	AS	ALi	A
SF	0,5																	
SGL	-1,0	-1,0																
SL	0,2	0,3	-1,0															
SFL	0,2	0,2	-1,0	0,1														
STFL	0,8	0,8	-1,0	0,7	0,7													
SFL	0,7	0,7	-1,0	0,6	0,6	0,7												
LS	0,6	0,5	-1,0	0,5	0,5	1,0	0,8											
LSF	0,3	0,2	-1,0	0,1	0,1	0,8	0,6	0,5										
Li	0,4	0,3	-1,0	0,2	0,2	0,9	0,7	0,6	0,1									
L	0,2	0,3	-1,0	0,1	0,2	0,8	0,7	0,6	0,1	0,1								
LLi	0,4	0,3	-1,0	0,2	0,3	0,9	0,8	0,7	0,2	0,1	0,1							
LSA	0,8	0,6	-1,0	0,6	0,5	0,9	0,9	0,8	0,5	0,6	0,5	0,5						
LA	1,3	1,2	-1,0	1,3	1,2	1,4	1,4	0,7	1,2	1,3	1,3	1,3	1,3					
LLiA	0,5	0,4	-1,0	0,3	0,3	0,7	0,8	0,7	0,3	0,3	0,2	0,2	0,3	1,3				
AS	-1,0	-1,0	-1,0	-1,0	-1,0	-1,0	-1,0	-1,0	-1,0	-1,0	-1,0	-1,0	-1,0	-1,0	-1,0	-1,0		
ALi	0,5	0,6	-1,0	0,4	0,5	0,6	0,8	0,9	0,5	0,6	0,4	0,5	0,5	1,4	0,2	-1,0		
A	0,6	0,5	-1,0	0,3	0,5	0,6	0,7	0,8	0,4	0,5	0,3	0,4	0,4	1,3	0,1	-1,0	0,1	
Alo	0,7	0,7	-1,0	0,5	0,6	0,7	0,9	1,0	0,6	0,8	0,5	0,7	0,5	1,4	0,3	-1,0	0,1	0,1
Moyenne : 0,26      Médiane : 0,46      Écart-type : 0,7																		

On peut remarquer que toutes les classes ne se séparent pas bien les unes des autres, à l'exception du loam argileux, avec les 6 indices spectraux utilisés. Ce scénario se répète pour toutes les images. Des classifications avec ce paramètre ont d'ailleurs révélé des taux de précision en bas de 20%.

Les classes sous-représentées ont donc été enlevées et les sols organiques non considérés pour se concentrer uniquement sur la granulométrie, cela améliore légèrement la détection par le logiciel



mais le problème est toujours présent, comme on peut le voir avec la distance de Bhattacharyya et la divergence transformée suivantes (Tableaux 6 et 7).

Tableau 6. Distance de Bhattacharyya avec les classes suffisamment représentées (sauf organique)  
avec l'image #9 de l'étude

	SL	SFL	LS	LSF	L	LLi	LA	LLiA	Ali
SFL	0,3								
LS	0,6	0,3							
LSF	0,1	0,2	0,6						
L	0,1	0,4	0,7	0,1					
LLi	0,1	0,4	0,7	0,1	0,0				
LA	1,0	0,8	0,3	1,0	1,1	1,1			
LLiA	0,2	0,5	0,8	0,2	0,1	0,1	1,1		
Ali	0,4	0,8	1,0	0,5	0,4	0,5	1,2	0,2	
A	0,3	0,7	0,9	0,4	0,3	0,4	1,1	0,1	0,0
Moyenne : 0,5		Médiane : 0,42					Écart-type : 0,35		

Tableau 7. Divergence transformée avec les classes suffisamment représentées (sauf organique)  
avec l'image #9 de l'étude

	SL	SFL	LS	LSF	L	LLi	LA	LLiA	Ali
SFL	0,3								
LS	1,1	0,3							
LSF	0,1	0,3	1,1						
L	0,1	0,5	1,4	0,1					
LLi	0,1	0,5	1,4	0,1	0,0				
LA	1,9	1,3	0,3	1,9	2,0	2,0			
LLiA	0,2	0,8	1,5	0,3	0,1	0,1	2,0		
Ali	0,5	1,0	1,7	0,6	0,4	0,5	2,0	0,2	
A	0,4	0,9	1,6	0,4	0,3	0,4	2,0	0,1	0,1
Moyenne : 0,78		Médiane : 0,50					Écart-type : 0,69		

Bien que cela propose une diversité de résultats bien inférieure, les classes ont été agrégées et réduites au maximum, en 3 classes (tableau 8).

Tableau 8. Agrégation des classes en argile, loam et sable

ARGILE	A
ARGILE	ALi
LOAM	L
LOAM	LA
LOAM	LLi
LOAM	LLiA
LOAM	LS
LOAM	LSF
SABLE	SL
SABLE	SFL

Ici encore, la séparation des classes est loin d'être parfaite malheureusement, mais comme l'agrégation s'est déjà faite au maximum, la classification des images s'est faite selon celle-ci. Le Tableau 9 suivant montre comment la séparation finale des classes est reconnue dans *Geomatica*.

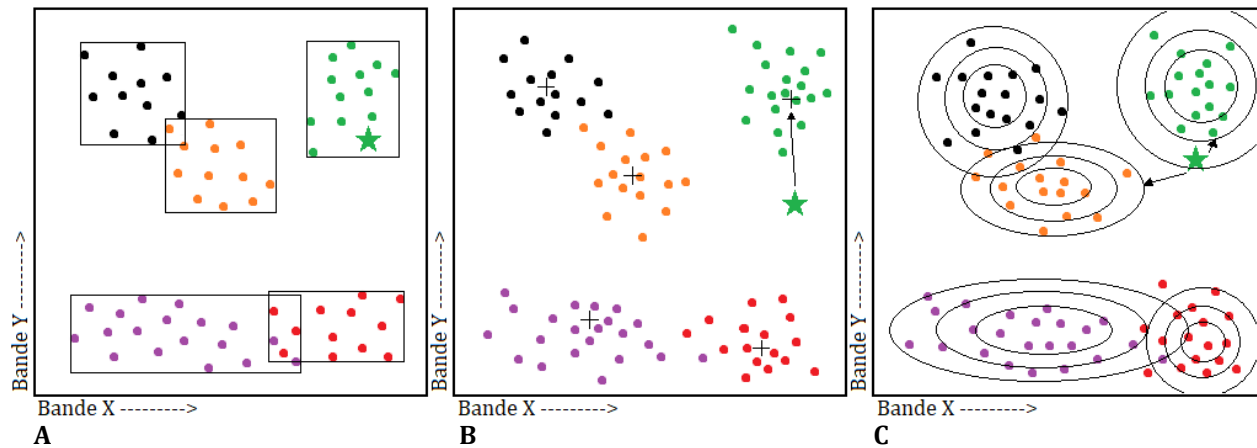
Tableau 9. Distance de Bhattacharyya et divergence transformée avec classes agrégées (image #9)

<b>A</b>	Sable	Loam	<b>B</b>	Sable	Loam
Loam	0,226		Loam	0,268	
Argile	0,724	0,953	Argile	1,027	1,706

Où : **A** = Distance de Bhattacharyya, **B** = Divergence transformée

On peut remarquer que le loam et le sable ainsi que le sable et l'argile se séparent mal et que l'argile et le loam ont la séparation la moins problématique. Dans tous les cas, ce n'est pas suffisant (trop loin de 1,9), mais on garde ici les classes principales.

*Geomatica* offre trois méthodes de classification (Figure 11), soit celle par parallélépipèdes (classification du pixel selon sa position par rapport à des rectangles de classe), par distance minimale (classification du pixel selon sa position par rapport à la moyenne de classe la plus proche) et par maximum de vraisemblance (classification du pixel par rapport à la classe avec laquelle la probabilité qu'il s'y situe est la plus grande). Il est recommandé d'utiliser la méthode du maximum de vraisemblance, qui nécessite plus de temps, mais donne généralement de meilleurs résultats (Duarte et al., 2018).



A : parallélipèdes; B : distance minimale; C : maximum de vraisemblance; étoile verte : exemple de pixel à classer; couleurs de points : différentes classes

Figure 11. Méthodes de classifications sur PCI

Sans compter sa supériorité potentielle (Duarte et al., 2018), la méthode par maximum de vraisemblance s'est trouvée être la plus efficace; les matrices de confusion post-classification initiales révélaient des résultats supérieurs. C'est donc ainsi que toutes les images ont été classifiées.

Cela s'est fait, rappelons-le, sur *PCI*, en utilisant les points *in situ* (transformées sites d'entraînements de 10 x 10 mètres) comme données d'élaboration des caractéristiques de classes et les indices spectraux (Tableau 1), calculés à partir des images satellites, comme données en intrants.

## 4. RÉSULTATS

### 4.1. CARTOGRAPHIE

La cartographie a été entièrement effectuée sur QGIS. On peut au premier coup d'œil en retenir que l'argile semble majoritairement « détectée » sur l'ensemble de la zone, aux dépens du sable mais surtout du loam, qui est très rare. Voir la Figure 12 suivante pour un exemple, ainsi que les Figures 15 à 25 en annexe pour l'ensemble des cartes.

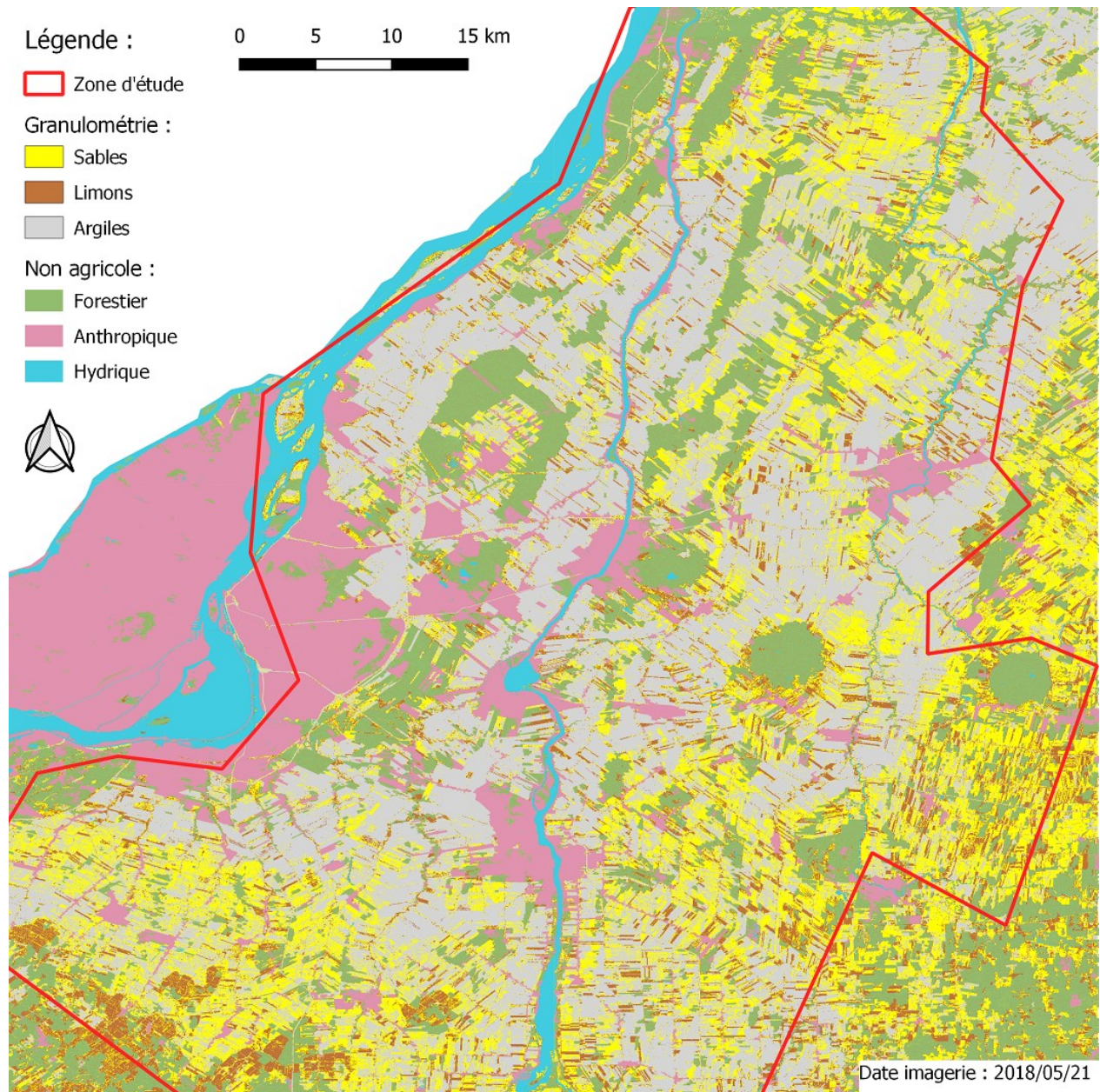


Figure 12. Granulométrie de la zone d'étude, classifiée avec l'image #9, du 2018/05/21

## 4.2. STATISTIQUES DES CLASSIFICATIONS

L'information statistique neutre la plus pertinente peut être la disposition des classes par couverture, coupée selon la zone d'étude, au Tableau 10 ci-dessous :

Tableau 10. Disposition des classes

Couverture	Classes	Pixels	%
21/04/2016	Sables	9 975 507	46,3
	Loams	3 592 350	16,7
	Argiles	7 988 803	37,1
23/04/2018	Sables	12 001 795	29,0
	Loams	9 183 484	22,2
	Argiles	20 271 383	48,9
08/05/2018	Sables	6 387 888	15,4
	Loams	10 243 632	24,7
	Argiles	24 824 031	59,9
13/05/2018	Sables	10 872 704	26,2
	Loams	18 706 294	45,1
	Argiles	11 920 287	28,7
16/05/2018	Sables	13 195 118	27,2
	Loams	10 245 608	21,1
	Argiles	25 057 163	51,7
21/05/2018	Sables	13 547 754	27,9
	Loams	7 595 539	15,7
	Argiles	27 354 596	56,4
31/05/2018	Sables	16 769 338	34,6
	Loams	7 595 273	15,7
	Argiles	24 126 986	49,8
10/06/2018	Sables	13 385 679	27,6
	Loams	4 198 373	8,7
	Argiles	30 913 837	63,7
Agrégat des couvertures	Sables	15 749 014	31,7
	Loams	9 247 476	18,6
	Argiles	24 650 705	49,7
Moyenne des %	Sables		29,3
	Loams		21,2
	Argiles		49,5

Où : Agrégat des couvertures = Matrice de l'ensemble des couverture agrégées (moyenne), Moyenne des % = Moyenne des pourcentages de chaque couverture, excluant l'agrégat.



Puisqu'il est principalement ici question d'établir si la méthode fonctionne ou non, les rapports de classification sont plus parlants que les cartes ou que la proportion des classes. Lorsque l'on regarde les matrices de confusion, soit la comparaison des images classifiées aux sites d'entraînement (points au sol), on peut voir que les résultats sont généralement peu prometteurs lorsqu'une seule image est utilisée (Tableau 11).

Tableau 11. Statistiques principales des matrices de confusion

Statistique	Valeur
Précision moyenne totale par classe	46,22 %
Précision moyenne totale réelle	26,89 %
Coefficient de Kappa moyen	0,161

La précision moyenne par classe est établie selon la précision de chaque classe dans sa propre identification, sans compter l'aire des sites d'entraînement. La précision moyenne réelle compte l'aire des sites d'entraînement. Le coefficient de Kappa (entre 0 et 1) signifie ici, selon l'étude de Landis et Koch (1977), un degré mauvais d'accord avec 0,161 (Voir Tableau 15).

Les matrices de confusion pour chaque couverture sont au Tableau 13, à la page suivante. Le Tableau 12 suivant montre la matrice de confusion pour trois couvertures : la première (**A**) étant basée sur l'agrégat des indices spectraux (par moyenne) de toutes les couvertures; la deuxième (**B**) étant basée sur l'agrégat des indices spectraux (par moyenne) des quatre couvertures représentant la grande majorité de la zone d'étude et donc des points de référence (couvertures du 18/05/16, 18/05/21, 18/05/31 et 18/06/10); et la troisième (**C**) étant l'agrégat post-classification (par mode) des 3 meilleures classifications (couvertures du 18/05/16, 18/05/21 et 18/06/10).

Tableau 12. Matrices de confusion (%) après classification par maximum de vraisemblance pour l'ensemble des couvertures agrégées et l'agrégat des meilleures classifications

Agrégat A				Agrégat B				Agrégat C			
	Sables	Loams	Argiles		Sables	Loams	Argiles		Sables	Loams	Argiles
Sables	46,92	4,05	49,03	Sables	35,69	1,22	63,10	Sables	43,28	1,11	55,61
Loams	37,90	3,56	58,54	Loams	27,73	1,45	70,83	Loams	31,45	1,27	67,28
Argiles	7,41	2,01	90,58	Argiles	10,65	0,97	88,38	Argiles	5,75	0,68	93,57
Précision moy. (classes) :	47,02			Précision moy. (classes) :	41,89			Précision moy. (classes) :	46,04		
Précision moy. réelle :	29,91			Précision moy. réelle :	27,37			Précision moy. réelle :	29,15		
Coefficient de kappa :	0,318			Coefficient de kappa :	0,290			Coefficient de kappa :	0,086		

La première (A) semble être la plus prometteuse, avec un coefficient de kappa de 0,318.

Tableau 13. Matrices de confusion (%) après classifications par maximum de vraisemblance (par année de couverture)

21/04/2016				16/05/2018			
	Sables	Loams	Argiles		Sables	Loams	Argiles
Sables	88,65	10,04	1,31	Sables	46,03	2,91	51,06
Loams	76,01	16,91	7,07	Loams	35,00	3,48	61,52
Argiles	48,61	33,61	17,78	Argiles	6,34	1,97	91,69
Précision moyenne (classes) :			41,11	Précision moyenne (classes) :			47,07
Précision moyenne réelle :			17,66	Précision moyenne réelle :			30,29
Coefficient de Kappa :			-0,003	Coefficient de Kappa :			0,321
23/04/2018				21/05/2018			
	Sables	Loams	Argiles		Sables	Loams	Argiles
Sables	56,54	0,95	42,52	Sables	47,19	2,20	50,61
Loams	35,63	1,16	63,21	Loams	33,38	2,48	64,14
Argiles	9,60	0,84	89,57	Argiles	5,91	1,53	92,56
Précision moyenne (classes) :			49,09	Précision moyenne (classes) :			47,41
Précision moyenne réelle :			24,76	Précision moyenne réelle :			29,94
Coefficient de Kappa :			0,084	Coefficient de Kappa :			0,318
08/05/2018				31/05/2018			
	Sables	Loams	Argiles		Sables	Loams	Argiles
Sables	37,32	3,83	58,85	Sables	55,88	1,95	42,16
Loams	16,11	4,67	79,22	Loams	44,14	2,26	53,60
Argiles	3,03	1,74	95,23	Argiles	15,14	1,20	83,66
Précision moyenne (classes) :			45,74	Précision moyenne (classes) :			47,27
Précision moyenne réelle :			28,04	Précision moyenne réelle :			23,97
Coefficient de Kappa :			0,071	Coefficient de Kappa :			0,082
13/05/2018				10/06/2018			
	Sables	Loams	Argiles		Sables	Loams	Argiles
Sables	61,31	11,11	25,57	Sables	34,08	1,23	64,69
Loams	43,53	18,31	38,15	Loams	26,09	1,31	72,59
Argiles	19,63	8,05	72,32	Argiles	10,19	0,90	88,90
Précision moyenne (classes) :			50,65	Précision moyenne (classes) :			41,43
Précision moyenne réelle :			33,11	Précision moyenne réelle :			27,33
Coefficient de Kappa :			0,125	Coefficient de Kappa :			0,289

Ces statistiques plutôt basses s'expliquent par une différenciation des classes ultimement faible, le Tableau 14 révèle la différenciation des 3 classes pour chaque couverture.

Tableau 14. Distance de Bhattacharyya (**DB**) et divergence transformée (**DT**) pour chaque couverture et les agrégats

21/04/16 :						21/05/18 :					
<b>DB</b>	Sable	Loam	<b>DT</b>	Sable	Loam	<b>DB</b>	Sable	Loam	<b>DT</b>	Sable	Loam
Loam	0,500		Loam	0,602		Loam	0,226		Loam	0,268	
Argile	1,070	0,505	Argile	1,501	0,630	Argile	0,724	0,953	Argile	1,027	1,706
23/04/18 :						31/05/18 :					
<b>DB</b>	Sable	Loam	<b>DT</b>	Sable	Loam	<b>DB</b>	Sable	Loam	<b>DT</b>	Sable	Loam
Loam	0,501		Loam	0,746		Loam	0,858		Loam	1,707	
Argile	0,466	0,751	Argile	0,571	1,264	Argile	0,343	1,067	Argile	0,373	1,953
08/05/18 :						10/06/18 :					
<b>DB</b>	Sable	Loam	<b>DT</b>	Sable	Loam	<b>DB</b>	Sable	Loam	<b>DT</b>	Sable	Loam
Loam	0,702		Loam	0,675		Loam	0,509		Loam	0,81	
Argile	0,801	0,555	Argile	1,003	0,977	Argile	0,479	1,026	Argile	0,581	1,918
13/05/18 :						Agrégat A :					
<b>DB</b>	Sable	Loam	<b>DT</b>	Sable	Loam	<b>DB</b>	Sable	Loam	<b>DT</b>	Sable	Loam
Loam	0,225		Loam	0,238		Loam	1,468		Loam	2,000	
Argile	0,739	0,445	Argile	0,938	0,493	Argile	0,594	1,629	Argile	0,827	2,000
16/05/18 :						Agrégat B :					
<b>DB</b>	Sable	Loam	<b>DT</b>	Sable	Loam	<b>DB</b>	Sable	Loam	<b>DT</b>	Sable	Loam
Loam	0,851		Loam	0,620		Loam	0,509		Loam	0,816	
Argile	0,499	0,420	Argile	0,603	1,490	Argile	0,453	1,015	Argile	0,542	1,903

Où : Agrégat A et B font référence aux classifications du Tableau 12

L'agrégat A du précédent démontre la meilleure séparation de classes, bien que le sable et l'argile y sont toujours mal séparables. Ceux-ci sont tout particulièrement mal séparables dans les couvertures du 13/05/18 et 21/05/18, avec des distance de Bhattacharyya de 0,225 et 0,226 respectivement.

Tableau 15. Signification du coefficient de Kappa

K	Accord	K	Accord
< 0	Négatif	0,41- 0,60	Moyen
0,0 – 0,20	Très faible	0,61 – 0,80	Fort
0,21 – 0,40	Faible	0,81 – 1,00	Très fort



## 5. DISCUSSION

### 5.1. DISPOSITION GÉOGRAPHIQUE

Bien que la classification par la méthode du maximum de vraisemblance d'une seule image Sentinel-2 ne semble pas être sur la voie de proposer des résultats viables, les cartographies démontrent une certaine homogénéité dans leur classification, excluant la couverture du 2016-04-23, qui diffère beaucoup des autres, possiblement à cause du nombre très limité de points de référence. Des points communs que l'on retrouve sont :

- la prédominance de l'argile, surtout au centre;
- la très faible quantité de loam;
- la tendances de certaines zones, comme autour du Mont Rougemont (Figure 13), à être classées soit en loam, soit en sable, mais peu en argile;
- les zones autour des endroits végétalisés sont portées à être considérées comme du sable;
- les zones classées comme du sable sont aussi portées à être concentrées au nord, sud et à l'est des couvertures, formant une couronne autour du centre.

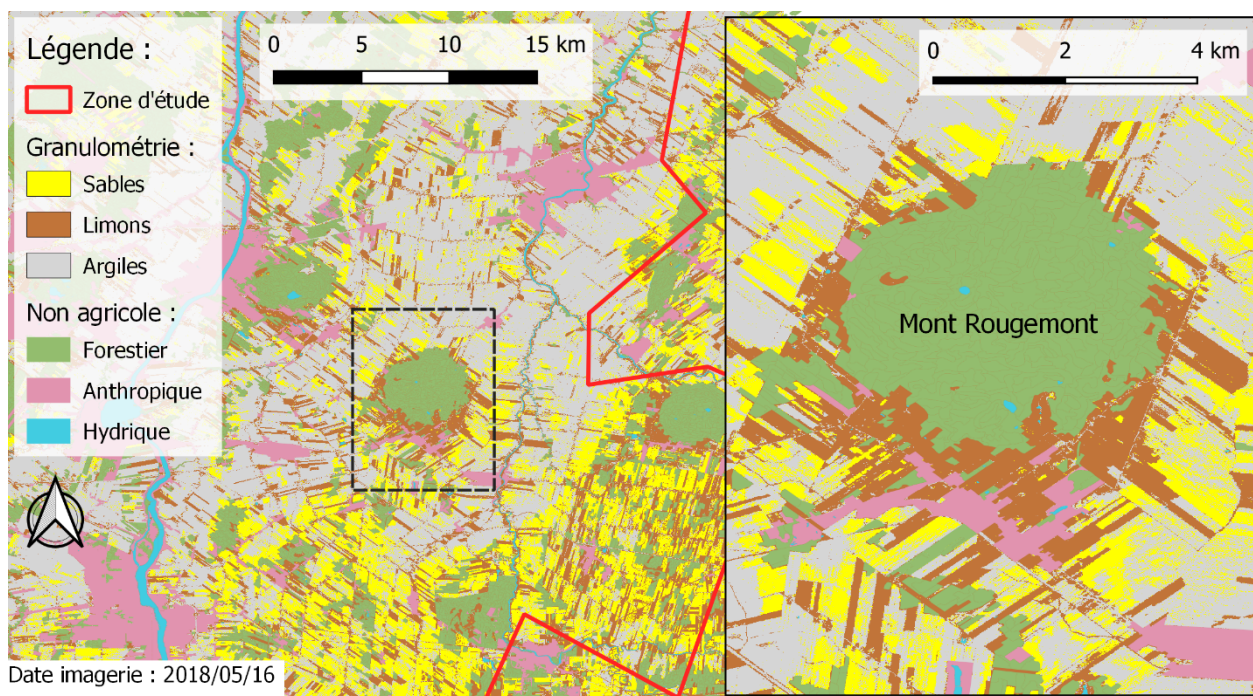


Figure 13. Agrandissement sur le Mont Rougemont, autour duquel le loam est souvent « repéré »

Ces particularités semblent pointer vers un lien que les matrices de confusion et les différenciations et proportions de classes ne peuvent révéler à elles seules, il serait potentiellement intéressant d'explorer cette avenue.

## 5.2. STATISTIQUES

Les résultats peuvent être séparés en 3 niveaux de signification, (1) le rejet automatique, (2) l'ultimement non significatif et (3) le légèrement prometteur.

1. Le rejet automatique concerne la couverture du 21/04/2016, qui est la moins significative du lot. Avec un coefficient de Kappa de -0,003, il va sans dire qu'elle peut être rejetée entièrement. Cette situation n'est pas particulièrement surprenante, considérant qu'elle ne représente qu'une scène à l'extrême sud de la zone, sans grande répartition géographique donc, et qu'elle comporte bien moins de points de référence.
2. L'ultimement non significatif concerne cinq couvertures : celles du 23/04/2018, 08/05/2018, 13/05/2018 et 31/05/2018 ainsi que l'agrégat C du Tableau 12 (en 4.2), avec des coefficients de Kappa de 0,084, 0,071, 0,125, 0,082 et 0,086 respectivement. Pour les trois premières, cela est sans doute lié au fait que leur disposition géographique est inférieure, ne représentant qu'environ la moitié de la zone d'étude, et que par extension les points de référence sont en nombre restreint.
3. Le légèrement prometteur comprend cinq couvertures : celles du 16/05/2018, 21/05/2018 et du 10/06/2018 ainsi que les agrégats A et B du Tableau 12 (en 4.2). Avec des coefficients de Kappa de 0,321, 0,318, 0,289, 0,318 et 0,290 respectivement, ils peuvent être considérés comme potentiellement significatifs d'un lien faible (Landis et Koch, 1977). Ces couvertures sont, sans surprise, celles qui représentent les plus grandes étendue géographique et nombre de points de référence. La séparation de leurs classes s'est donc faite de façon plus efficace (Tableau 13 en 4.2). La seule exception à cela est la couverture du 31/05/2018 qui, malgré qu'elle possède la même étendue géographique, montre des résultats non révélateurs. Il est possible que cela soit dû à son nombre légèrement inférieur en points de référence (malgré la large étendue) puisqu'elle comportait une certaine couverture nuageuse qu'il a fallu discriminer.

Ces 3 niveaux de signification sont donc largement influencés par le nombre de points de référence et l'étendue géographique, il est suggéré de retourner à la Figure 7 ainsi qu'au Tableau 4 à la section 3.2.2 pour bien comprendre ce point. Il aurait pu être assumé que les classifications des agrégats proposeraient des résultats notablement supérieurs à ceux des couvertures uniques, elles en proposent cependant des semblables.

L'argile est une classe qui s'est montrée presque toujours « détectée » en majorité, le sable généralement en deuxième et le loam est rarement révélé, bien qu'il représente une partie importante

des points de référence. Il y a donc des éléments qui brouillent beaucoup la distinction des classes, encore une fois, et des travaux supplémentaires seraient nécessaires. Il est important de noter que bien que presque tous les paramètres utilisés pour la correction atmosphérique étaient suffisamment adaptés, l'un ne l'était pas forcément, soit la profondeur optique des aérosols, qui était une approximation. Cela influe sur les indices spectraux finaux utilisés pour la classification, bien que celle-ci soit probablement davantage liée à la définition des classes.

La présence de végétation est certainement un facteur négatif dans les classifications effectuées. Comme expliqué en 3.2.2, la discrimination de la végétation s'est faite avec un NDVI inférieur à 0,35 afin de conserver un nombre important de points, tout en éliminant la végétation dense. Cela a laissé plusieurs endroits où étaient possiblement présentes des traces de végétations ou même des plantes de tailles non négligeables. La végétation morte influe moins sur le NDVI mais affecte toutes les réflectances en générale, et donc les indices spectraux utilisés pour le sol. Une autre limite importante concerne la résolution spatiale des images (10 m) qui ne permet de discriminer les sols à des échelles intra-parcellaire fine. Pour potentiellement améliorer les résultats, il serait intéressant de répéter l'étude avec plusieurs niveaux de discrimination selon le NDVI, avec à chaque fois un compromis différent entre le nombre de points de comparaison et la présence de végétation. Un tel travail pourrait d'ailleurs très bien être effectué puisque le téléchargement et la correction atmosphérique des données se sont effectués de manière semi-automatique. Des images satellites à très haute résolution spatiale (autour de 2 m) ou même de l'imagerie prise par drone (résolution centimétrique) auraient donné de meilleurs résultats, mais aux dépens évidemment d'un coût élevé et d'une étendue spatiale bien inférieure.

Comme dernier test, une classification additionnelle a été effectuée en ne prenant compte que de certains points; ceux avec une séparabilité plus raisonnable encore (Tableau 15). Cela aux dépens de la classe de sable, non considérée. Cette classification est donc sévèrement limitée et servirait principalement à établir si un point de l'image a plus de chance d'être argileux que loameux (Tableau 16), sans prendre compte la possibilité d'être quelque chose d'autre entièrement. Seule la couverture avec les meilleurs résultats a été utilisée, soit celle du 16/05/2018.

Tableau 16. Classes conservées pour la classification à deux classes (16/05/2018)

	SL	SFL	LS	LSF	L	Lli	LA	LLiA	Ali
SFL	0,3								
LS	0,6	0,3							
LSF	0,1	0,2	0,6						
L	0,1	0,4	0,7	0,1					
Lli	0,1	0,4	0,7	0,1	0,0				
LA	1,0	0,8	0,3	1,0	1,1	1,1			
LLiA	0,2	0,5	0,8	0,2	0,1	0,1	1,1		
Ali	0,4	0,8	1,0	0,5	0,4	0,5	1,2	0,2	
A	0,3	0,7	0,9	0,4	0,3	0,4	1,1	0,1	0,0

Où : Les cases en blanc concordent avec les classes supprimées

Tableau 17. Matrice de confusion de la classification à deux classes (16/05/2018)

	Loams	Argiles
Loams	16,59	83,41
Argiles	2,92	97,08

Précision moyenne (classes) : 56,84

Précision moyenne réelle : 60,61

Coefficient de Kappa : 0,147

On obtient alors une précision moyenne par classe de 57%, une précision moyenne réelle de 61% et un coefficient de Kappa de 0,147. Cette classification à deux possibilités est donc peu prometteuse. La carte réalisée à partir de cette classification est disponible en annexe à la Figure 23.

### 5.3. COMPARAISON AVEC D'AUTRES DONNÉES ET UN AUTRE CLASSIFICATEUR

Cette section montre, à titre comparatif, des résultats obtenus par un autre groupe de recherche pour la même problématique de ce projet. L'étude a été réalisée par l'équipe d'Observation de la Terre de la compagnie Effigis Géo-Solutions, mais en utilisant douze (12) images Landsat-5, -7 et -8 et un classificateur statistique basé sur une « Analyse factorielle discriminante », ou AFD (Effigis, 2019b). Les points in-situ retenus (1 901 seulement) sont tous en sol nu sur les 12 images. Les classes de sol sont également différentes de celles que nous avons considérées (Tableau 17). Dans ces conditions, les taux de succès sont plus élevés que ceux que nous avons obtenus dans notre étude : les taux de succès varient de 53 à 83%. Si on tolère les classifications dans les classes voisines (ex. : A classé A

vers L, L classé L vers S, etc.), les taux de succès vont de 79 à 94% pour les sols minéraux. Notons que l'équipe d'Effigis n'a pas obtenu de résultats concluants avec les images Sentinel-2, tout comme notre étude.

Tableau 18. Résultats d'identification des textures de sol à partir de 12 images Landsat par la méthode AFD (analyse factorielle discriminante), étude de l'équipe d'OT d'Effigis

	Argile	A. vers L.	Loam	L. vers S.	Sable	Organique	Total	% correct	% avec tolérance
A	409	54	20	5	0	3	491	0,83	0,94
Argile vers Loam	73	199	54	35	3	13	377	0,53	0,86
Loam	26	62	217	99	5	14	423	0,51	0,89
Loam vers Sable	14	24	92	268	6	22	426	0,63	0,86
Sable	0	1	4	3	11	0	19	0,58	0,79
Organique	0	20	20	21	1	103	165	0,62	0,63
Total	522	360	407	431	26	155	1901	0,63	0,87

Source : Effigis, 2019b

L'utilisation d'un classificateur de type réseau de neurones profond a donné des résultats encore meilleurs, selon l'équipe d'Effigis. Ces résultats ne sont pas montrés ici, ils feront l'objet de diffusion dans les mois à venir, selon Effigis.

## 6. CONCLUSION

Cette étude a montré les limites de l'utilisation d'une seule image Sentinel-2 avec un classificateur de type « maximum de vraisemblance » pour la cartographie de la texture des sols agricoles. Les classes de texture se séparent mal malgré les nombreuses tentatives et divers tests et la végétation (bien que très faible) a fort probablement joué un rôle plus important que ce qui a été prévu au début, la plupart des statistiques et cartographies créées sont donc à considérer avec beaucoup de réserve.

La comparaison à l'utilisation de plus d'images (12 scènes Landsat, selon une étude similaire réalisée par Effigis) montre la possibilité d'améliorer considérablement les résultats. Ceci montre l'importance d'utiliser des scripts automatiques pour la recherche, la récupération et le traitement des images. La méthode d'acquisition des images à l'aide d'un script *Java* sur *Google Earth Engine* s'est révélée très intéressante, elle permet de faire des sélections selon des caractéristiques que l'on peut adapter au besoin, dépendamment de la géographie, de l'imagerie nécessaire, des bandes spectrales à utiliser, etc.

La méthode de correction atmosphérique, bien qu'elle exige la connaissance de la profondeur optique des aérosols qui est difficile à évaluer, utilise un code *Python* sur *GRASS Gis* qui permet de corriger de façon semi-automatique un nombre important d'images Sentinel-2 (et potentiellement des images d'autre provenance), peu importe leur taille (en considérant que l'ordinateur utilisé supporte la charge mémoire). Le script peut être modifié pour combler différents modèles atmosphériques, d'aérosols, altitude, etc. Une fois le code lancé, il suffit d'attendre que le logiciel finisse son traitement, ce qui permet le multitâches, un aspect intéressant aux niveaux professionnel ou scientifique, où le temps peut souvent manquer.

Ces deux scripts ont permis, par la facilité améliorée d'acquisition ainsi que la simplification notable de l'étape de correction atmosphérique, la comparaison multi-imageries, qui aurait été complexifiée et possiblement faussée si les précédentes étapes avaient été effectuées à la main. Dans le contexte grandissant du *Big Data*, et plus précisément dans un contexte d'utilisation de plusieurs images représentant la même localisation comme c'est le cas ici, cela constitue un aspect technique intéressant.

Si le travail était à refaire, ou si quelqu'un voulait en effectuer une recherche semblable, il serait pertinent d'abord de se concentrer sur l'établissement des classes et de s'assurer que celles-ci peuvent facilement être séparées. L'utilisation d'algorithmes de classification avancées (ex.; un réseau de neurones profond) est également recommandée. Ensuite, il faudrait reconsidérer la

discrimination de la couverture du sol par le NDVI et faire plus de tests, afin d'obtenir le meilleur compromis entre le nombre de points et la réduction de l'effet de la végétation. Une étude similaire à Effigis, qui a utilisé un nombre de points bien moins important, a donné des résultats supérieurs. Cela nécessite beaucoup de travail et de traitements, donc de temps, mais peut valoir la peine, d'autant plus que les étapes d'acquisition et de correction atmosphérique des images peuvent être, comme mentionné ci-haut, raccourcies et simplifiées.



## BIBLIOGRAPHIE

Agriculture et Agroalimentaire Canada (2013). « Banque de données des profils de sol de la Montérégie - Données morphologiques », Laboratoires de pédologie et d'agriculture de précision, Québec, Canada.

Agriculture et Agroalimentaire Canada (2014). « Texture du sol et qualité de l'eau », Gouvernement du Canada, [URL]. (<http://www.agr.gc.ca/fra/science-et-innovation/pratiques-agricoles/sol-et-terre/le-sol-et-l-eau/texture-du-sol-et-qualite-de-l-eau/?id=1197483793077>). Page visitée le 22 avril 2019.

Agriculture et Agroalimentaire Canada (2017). « SCAN, un outil d'agriculture de précision pour optimiser la gestion de l'azote aux champs », Agriculture et Agroalimentaire Canada, [URL]. (<https://www5.agr.gc.ca/fra/nouvelles/realisations-scientifiques-en-agriculture/scan-un-outil-dagriculture-de-precision-pour-optimiser-la-gestion-de-lazote-aux-champs/?id=1496858857623>). Page visitée le 23 février 2019.

Amazon (2019). « Earth », [URL]. (<https://aws.amazon.com/fr/earth/>). Page visitée le 23 février 2019.

Bouroubi, Y., Surprenant, J., Bugnet, P., Nguyen-Xuan, T., Tremblay, N., Grenon, L. et Vigneault, P. (2016). « Cartographie des sols agricoles à l'échelle du champ par imagerie satellitaire », présentation au colloque : Agriculture : les enjeux scientifiques, technologiques, et socio-économiques, au Département de Géographie de l'Université de Montréal, Mercredi le 23 mars 2016. [En ligne]. ([https://geographie.umontreal.ca/fileadmin/Documents/FAS/geographie/Documents/Nouvelles/5\\_Bouroubi\\_Yacine\\_CARTOGRAPHIE\\_DES\\_SOLS\\_AGRICOLES.pdf](https://geographie.umontreal.ca/fileadmin/Documents/FAS/geographie/Documents/Nouvelles/5_Bouroubi_Yacine_CARTOGRAPHIE_DES_SOLS_AGRICOLES.pdf)). Page consultée le 1er novembre 2018.

Bouroubi, Y., Cavayas, F., Tremblay, N. et Batita, W. (2018). « Ground reflectance retrieval on horizontal and inclined terrains using the software package REFLECT », *Remote Sensing*, 10, 1638, 34 p.

Capmourteres, V., Adams, J., Berg, A., Fraser, E., Swanton, C. et Anand, M. (2018). « Precision conservation meets precision agriculture: A case study from southern Ontario », *Agricultural Systems*, 167, p. 176-185.

Clément, P., Bonn, F. et Dubois, J.-M. M. (1983). « Télédétection et caractéristiques physiques des dépôts meubles dans un milieu anciennement englacé du sud du Québec », *Géographie physique et Quaternaire*, 37, 1, p. 107-117.

CNES (2018). « Sentinel-2 : Instruments », [URL]. (<https://sentinel2.cnes.fr/fr/instruments-2>). Page visitée le 28 novembre 2018.

Duarte, D. C., Zanetti, J., Gripp, J. et Medeiros, N. (2018). « Comparison of supervised classification methods of Maximum Likelihood, Minimum Distance, Parallelepiped and Neural Network in images of Unmanned Air Vehicle (UAV) in Viçosa – MG », *Revue Brésilienne de Cartographie*, 70, 2.

Effigis (2019a). « SCAN », [URL]. (<https://www.effigis.com/en/scan/>). Page visitée le 28 février 2019.

Effigis (2019b). ScanSol : Cartographie intelligente des sols agricoles par fouille de mégadonnées d'OT. Rapport de recherche interne, mars 2019.

- ESA (2019a). « Level 1 », [URL]. (<https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/processing-levels/level-1>). Page visitée le 12 avril 2019.
- ESA (2019b). « Sentinel-2 », [URL]. (<https://sentinel.esa.int/web/sentinel/missions/sentinel-2>). Page visitée le 5 mars 2019.
- Google (2019a). « Google Drive », [URL]. ([https://www.google.com/intl/fr\\_ALL/drive/](https://www.google.com/intl/fr_ALL/drive/)). Page visitée le 2 mars 2019.
- Google (2019b). « Google Earth Engine », [URL]. (<https://earthengine.google.com/>). Page visitée le 2 mars 2019.
- Gholizadeh, A., Boruvka, L., Saberioon, M. et Vašát, R. (2016). « A Memory-Based Learning Approach as Compared to Other Data Mining Algorithms for the Prediction of Soil Texture Using Diffuse Reflectance Spectra », *Remote Sensing*, 8, 341, 17 p.
- Gili, A., Álvarez, C., Bagnato, R., Noellemeyer, E. (2017). « Comparison of three methods for delineating management zones for site-specific crop management », *Computers and Electronics in Agriculture*, 139, p. 213-223.
- OMAFRA (2012). « Texture du sol », Ministère de l'Agriculture, de l'Alimentation et des Affaires Rurales de l'Ontario, [URL]. ([http://www.omafr.gov.on.ca/CropOp/fr/general\\_agronomics/soil\\_management/soil\\_texture.html](http://www.omafr.gov.on.ca/CropOp/fr/general_agronomics/soil_management/soil_texture.html)). Page visitée le 13 mai 2019.
- Houssa, R., Pion, J.-C. et Yédou, H. (1996). « Effects of granulometric and mineralogical composition on spectral reflectance of soils in a Sahelian area », *ISPRS Journal of Photogrammetry & Remote Sensing*, 51, p. 284-298.
- Huang, Y., Chen, Z.-X., Yu, T., Huang, X.-Z., et Gu, X.-F. (2018). « Agricultural remote sensing big data: Management and applications », *Journal of Integrative Agriculture*, 17, 9, p. 1915-1931.
- IRDA (2019). « Documents pédologiques disponibles à l'IRDA », [URL]. (<https://www.irda.qc.ca/fr/services/protection-ressources/sante-sols/information-sols/etudes-pedologiques/>). Page visitée le 25 février 2019.
- Kergomard, C. (1996). « Pratique des corrections atmosphériques en télédétection : utilisation du logiciel 5S-PC », *Cybergeo : European Journal of Geography*, 181, [URL]. (<http://journals.openedition.org/cybergeo/1679>). Page visitée le 25 janvier 2019.
- Landis, JR. et Koch, G. (1977). « The measurement of observer agreement for categorical data », *Biometrics*, 33, p.159-174.
- Lokers, R., Knapen, R., Janssen, S., van Randen, Y. et Jansen, K. (2016). « Analysis of Big Data technologies for use in agro-environmental science », *Environmental Modelling & Software*, 84, p. 494-504.
- MAPAQ (2014). « Portrait agroalimentaire : la Montérégie », Ministère de l'Agriculture, Pêcheries et Alimentation du Québec – Directions régionales de la Montérégie, secteurs est et ouest, [URL]. (<https://www.mapaq.gouv.qc.ca/fr/Pages/Details-Publication.aspx?docid=DDJ7DZ3RAA3J-202-12342>). Page visitée le 22 février 2019.

MAPAQ (2018). « Production agricole », Ministère de l'Agriculture, Pêcheries et Alimentation du Québec, [URL]. (<https://www.mapaq.gouv.qc.ca/fr/md/statistiques/Pages/production.aspx>). Page visitée le 22 février 2019.

Mulla, D. et Khosla, R. (2015). « Historical evolution and recent advances in precision farming », Ch. 1 dans: Lal, R. et Stewart, B. A., éd., *Soil Specific Farming: Precision Agriculture*, Taylor and Francis : Boca Raton, Floride, ÉU.

Ose, K., Corpetti, T. et Demagistri, L. (2016). « 2 : Multispectral Satellite Image Processing », chap. dans (Édition : Baghdadi, N. et Zribi, M.) *Optical Remote Sensing of Land Surface – Techniques and Methods*, ISTE Press : Elsevier, 388 pages. ISBN : 978-1-78548-102-4.

OSGeo4W (2019). « Grass 7.6 manuals : i.atcorr », [URL]. (<https://grass.osgeo.org/grass76/manuals/i.atcorr.html>). Page visitée le 22 mars 2019.

OSGeo4W (2019b). « QGIS », [URL]. (<https://www.qgis.org/fr/site/>). Page visitée le 23 février 2019.

PCI Geomatics (2018). « Geomatica II – Course guide », [URL]. (<https://www.pcigeomatics.com/pdf/TrainingGuide-Geomatica-2.pdf>). Page visitée le 2 octobre 2019.

PCI Geomatics (2019). « PCI Geomatics », [URL]. (<https://www.pcigeomatics.com/>). Page visitée le 30 février 2019.

Plant, R. E. (2001). « Site-specific management: the application of information technology to crop production », *Computer and Electronics in Agriculture*, 30, p. 9-29.

Radian Earth Foundation (2019). « Radian Earth », [URL]. (<https://www.radiant.earth/>). Page visitée le 21 février 2019.

Siegal, B. S. et Goetz, A. F. H. (1977). « Effect of Vegetation on Rock and Soil Type Discrimination », *Photogrammetric Engineering and Remote Sensing*, 43, 2, p. 191-196.

Sharma, A. R., Badarinath, K. V. S. et Roy, P. S. (2009). « Comparison of ground reflectance measurement with satellite derived atmospherically corrected reflectance: A case study over semi-arid landscape », *Advances in Space Research*, 43, p. 56-64.

Sola, I., García-Martín, A., Sandonís-Pozo, L., Álvarez-Mozos, J., Pérez-Cabello, F., González-Audícana, M. et Llovería, R. M. (2018). « Assessment of atmospheric correction methods for Sentinel-2 images in Mediterranean landscapes », *International Journal of Applied Earth Observation and Geoinformation*, 73, p. 63-76.

Tremblay, N., Y. Bouroubi, C. Bélec, R. Mullen, N. Kitchen, W. Thomason, S. Ebelhar, D. Mengel, B. Raun, D. Francis, E.D. Vories, and I. Ortiz-Monasterio. 2012. Corn Response to Nitrogen is Influenced by Soil Texture and Weather. *Agronomy Journal*, 104(6): 1658-1671.

Vermote, E. F., Tanré, D., Deuzé, J. L., German, M. et Morcrette, J.-J., (1997). « Second Simulation of the Satellite Signal in the Solar Spectrum, 6S : An Overview », *IEEE Transactions on Geoscience and Remote Sensing*, 35, 3, p. 675-686.

Vermote, E., Tanré, D., Deuzé, J. L., Herman, M. Morcrette, J.-J. et Kotchenova, S. Y. (2006). « Second Simulation of a Satellite Signal in the Solar Spectrum - Vector (6SV) », [URL]. (<http://6s.ltdri.org/pages/manual.html>). Page visitée le 22 avril 2019.

USGS (s.d.). « Landsat-8 », [URL]. (<https://landsat.usgs.gov/landsat-8>). Page visitée le 26 février 2019.

Zhao, W., Tamura, M. et Takahashi, H. (2001). « Atmospheric and spectral corrections for estimating surface albedo from satellite data using 6S code », *Remote Sensing of Environment*, 76, p. 202-212.

## ANNEXE

Tableau 19. Images Sentinel-2 utilisées

1	ID_IMG	S2A_OPER_MSI_L1C_DS_SGS_20160421T194244_S20160421T160034_N02.01
	ID_GEE	COPERNICUS/S2/20160421T160034_20160421T194244_T18TWQ
	Date	21/04/2016 16:00:34
2	ID_IMG	S2A_OPER_MSI_L1C_DS_SGS_20160421T194244_S20160421T160034_N02.01
	ID_GEE	COPERNICUS/S2/20160421T160034_20160421T194244_T18TWR
	Date	21/04/2016 16:00:34
3	ID_IMG	S2A_OPER_MSI_L1C_DS_SGS_20160421T194244_S20160421T160034_N02.01
	ID_GEE	COPERNICUS/S2/20160421T160034_20160421T194244_T18TXQ
	Date	21/04/2016 16:00:34
4	ID_IMG	S2B_OPER_MSI_L1C_DS_EPAE_20180423T191817_S20180423T154331_N02.06
	ID_GEE	COPERNICUS/S2/20180423T153909_20180423T154331_T18TXR
	Date	23/04/2018 15:43:31
5	ID_IMG	S2A_OPER_MSI_L1C_DS_SGS_20180508T223752_S20180508T154947_N02.06
	ID_GEE	COPERNICUS/S2/20180508T153911_20180508T154947_T18TXR
	Date	08/05/2018 15:49:47
6	ID_IMG	S2B_OPER_MSI_L1C_DS_MTI_20180513T185554_S20180513T154810_N02.06
	ID_GEE	COPERNICUS/S2/20180513T153809_20180513T154810_T18TXR
	Date	13/05/2018 15:48:10
7	ID_IMG	S2B_OPER_MSI_L1C_DS_MTI_20180513T185554_S20180513T154810_N02.06
	ID_GEE	COPERNICUS/S2/20180513T153809_20180513T154810_T18TXS
	Date	13/05/2018 15:48:10
8	ID_IMG	S2B_OPER_MSI_L1C_DS_MTI_20180516T190834_S20180516T155840_N02.06
	ID_GEE	COPERNICUS/S2/20180516T154809_20180516T155840_T18TXR
	Date	916/05/2018 15:58:40
9	ID_IMG	S2A_OPER_MSI_L1C_DS_SGS_20180521T193200_S20180521T155356_N02.06
	ID_GEE	COPERNICUS/S2/20180521T154911_20180521T155356_T18TXR

	Date	21/05/2018 15:53:56
10	ID_IMG	S2A_OPER_MSI_L1C_DS_SGS_20180531T193310_S20180531T155122_N02.06
	ID_GEE	COPERNICUS/S2/20180531T155021_20180531T155122_T18TXR
	Date	31/05/2018 15:51:22
11	ID_IMG	S2A_OPER_MSI_L1C_DS_SGS_20180610T193029_S20180610T155519_N02.06
	ID_GEE	COPERNICUS/S2/20180610T154901_20180610T155519_T18TXR
	Date	10/06/2018 15:55:19

Où : ID\_IMG = identifiant original des images, ID\_GEE = identifiant sur *Google Earth Engine*, Date = Moment d'acquisition en temps UTC

Certaines images ont le même identifiant (ID\_IMG) et temps d'acquisition, mais sont représentées par des données séparées sur *Google Earth Engine*, d'où l'ID\_GEE.

Tableau 20. Bandes captées par le MSI de Sentinel-2

Résolution (m)	Bande	Satellite Sentinel-2A		Satellite Sentinel-2B	
		LO (nm)	LB (nm)	LO (nm)	LB (nm)
10	2	496.6	98	492.1	98
	3	560.0	45	559	46
	4	664.5	38	665	39
	8	835.1	145	833	133
20	5	703.9	19	703.8	20
	6	740.2	18	739.1	18
	7	782.5	28	779.7	28
	8a	864.8	33	864	32
	11	1613.7	143	1610.4	141
	12	2202.4	242	2185.7	238
60	1	443.9	27	442.3	45
	9	945.0	26	943.2	27
	10	1373.5	75	1376.9	76

Où : LO = Longueur d'onde, LB = Longueur de bande

Source : CNES, 2018

## Document 1. Code *Java* permettant d'accéder aux images Sentinel-2, sur *Google Earth Engine*

<pre>//DÉCLARATION DES INTERVALLES DE TEMPS POUR CHAQUE ANNÉE (PRINTEMPS) var p2000 = ee.Filter.date('2000-04-15','2000-06-15'); var p2001 = ee.Filter.date('2001-04-15','2001-06-15'); var p2002 = ee.Filter.date('2002-04-15','2002-06-15'); var p2003 = ee.Filter.date('2003-04-15','2003-06-15'); var p2004 = ee.Filter.date('2004-04-15','2004-06-15'); var p2005 = ee.Filter.date('2005-04-15','2005-06-15'); var p2006 = ee.Filter.date('2006-04-15','2006-06-15'); var p2007 = ee.Filter.date('2007-04-15','2007-06-15'); var p2008 = ee.Filter.date('2008-04-15','2008-06-15'); var p2009 = ee.Filter.date('2009-04-15','2009-06-15'); var p2010 = ee.Filter.date('2010-04-15','2010-06-15'); var p2011 = ee.Filter.date('2011-04-15','2011-06-15'); var p2012 = ee.Filter.date('2012-04-15','2012-06-15'); var p2013 = ee.Filter.date('2013-04-15','2013-06-15'); var p2014 = ee.Filter.date('2014-04-15','2014-06-15'); var p2015 = ee.Filter.date('2015-04-15','2015-06-15'); var p2016 = ee.Filter.date('2016-04-15','2016-06-15'); var p2017 = ee.Filter.date('2017-04-15','2017-06-15'); var p2018 = ee.Filter.date('2018-04-15','2018-06-15');  //AGRÉGAT DE L'ENSEMBLE DES PRINTEMPS var printemps = ee.Filter.or(p2000, p2001, p2002, p2003, p2004, p2005, p2006, p2007, p2008, p2009, p2010, p2011, p2012, p2013, p2014, p2015, p2016, p2017, p2018);  //CENTRAGE CARTE SUR MONTÉRÉGIE Map.setCenter(-73.12, 45.45, 10);  //POLYGONE MONTÉRÉGIE var monteregie = ee.Geometry.Polygon([ [[-72.9506313713033, 46.0959632130833],[-72.9625186074864, 46.0305834140759],[-72.8876290195326, 45.9711472331602],[- 72.8923839140058, 45.9342968009924],[-72.8246266677618, 45.8594072130385],[-72.8579109290747, 45.7880837959396],[- 72.8840628486776, 45.6430595145051],[-72.8519673109831, 45.605020358719],[-72.9368124592403, 45.5322110370972],[- 72.9375554115018, 45.4802043787959],[-72.8507785873648, 45.4932803385974],[-72.7958001200177, 45.4701745232663],[- 72.8721756124945, 45.2531581676976],[-72.9839156326161, 45.3137830722317],[-73.0837684165546, 45.0950579264617],[- 73.3084371804162, 45.0807932430419],[-73.3500425070573, 44.9963938661415],[-73.5568804166441, 45.0011487606148],[- 73.5295397734229, 45.0867368611335],[-73.7565859845211, 45.2555356149343],[-73.6816963965672, 45.3803515948574],[- 73.6139391503233, 45.3946162782772],[-73.5271623261862, 45.3839177657123],[-73.4631198412495, 45.4580644014048],[- 73.5033878538199, 45.5646037556963],[-73.4926893412551, 45.6977408009476],[-73.2454348286455, 45.8736718964583],[- 73.1859986477297, 46.0210736251294],[-72.9506313713033, 46.0959632130833]] ]);</pre>	<pre>//POLYGONE DE LA MONTÉRÉGIE POUR VISUALISATION Map.addLayer(monteregie, {color: 'red'}, 'MONTÉRÉGIE');  //ACQUISITION SENTINEL-2 var s2 = ee.ImageCollection("COPERNICUS/S2") .filterBounds(monteregie) .filterMetadata('CLOUD_COVERAGE_ASSESSMENT', 'less_than', 8) .filter(printemps);  //CRÉATION DE LA LISTE DES IMAGES TROUVÉES var liste = s2.toList(500); var count_list = liste.size().getInfo(); print('Images trouvées, cliquer pour étendre:',liste)  // PARCOURS DE LA LISTE D'IMAGES ET UTILISATION D'EXPORT.TODRIVE for (var i = 0; i &lt; count_list; i++) { var img = ee.Image(liste.get(i)); var id = img.id().getInfo();  var gen_time_info = ee.String(img.get('GENERATION_TIME')); var gen_time = gen_time_info.getInfo();  var porteur_info = ee.String(img.get('SPACECRAFT_NAME')); var porteur_info_complet = porteur_info.getInfo(); var porteur = porteur_info_complet.substr(-1);  // Pour changer le nombre de bandes voulues (5 ou 13), mettre la ligne indésirée en '/' : var image = img.select(['B4', 'B3', 'B2', 'B8', 'B11']); // var image = img.select(['B1', 'B2', 'B3', 'B4', 'B5', 'B6', 'B7', 'B8', 'B8A', 'B9', 'B10', 'B11', 'B12']);  Map.addLayer(image, {bands: ['B4', 'B3', 'B2'], min: 200, max: 2200}, (gen_time+porteur));  Export.image.toDrive({ image: image, description: (gen_time+porteur), folder : "Sentinel-2", crs : 'EPSG:4326', maxPixels : 900000000, region : monteregie, scale: 10 }) }</pre>
---	---



## Document 2. Code Python permettant d'effectuer les corrections atmosphériques sur GRASS Gis

```
#-----
# This python code uses raster data in a complementary way with a javascript code on
# Google Earth Engine. The latter one downloads Sentinel-2 data (only the images) that has as file
# names the original ID of the Sentinel-2 product.
# The scenes are also downloaded in the WGS-84 (EPSG:4326) projection. The GRASSS Gis project
# should therefore also be created in this projection.
#-----
#!/usr/bin/env python
#-*- coding: utf-8 -*-
import datetime
import os
import sys
import grass
import grass.script as grass
from grass.pygrass.modules.shortcuts import imagery as i

#-----
#          USER PARAMETERS
#-----

# This section is, in theory, the only one that has to be modified by the user:

# 1. input : . . . . . Folder containing the images.
# 2. Suffixe : . . . . . Files format (has to be raster data)
# 3. Modele_atmospherique : . See : https://grass.osgeo.org/grass76/manuals/i.atcorr.html
# 4. Modele_aerosols : . . . . . See : https://grass.osgeo.org/grass76/manuals/i.atcorr.html
# 5. Nombre_de_bandes : . . . . . For now, 5 (R, V, B, NIR, SWIR-1)
# 6. Visibilite . . . . . En KM, aerosols concentration.
# 7. Elevation_moyenne . . . . . En M, mean elevation of the scene's location.

input = r'C:\geodata\Essai\grass\in'
output = r'C:\geodata\Essai\grass\out'
Suffixe = '.tif'
Modele_atmospherique = 2
Modele_aerosols = 3
Nombre_de_bandes = 5
Visibilite = 90
Elevation_moyenne = 35

#OPTIONAL :
#Delete 6S parameters .txt files after completion? If so, type 'oui' or 'yes'. Leave blank otherwise.
Suppr = ''
#Use a shorter name for the exported corrected images? (date + sat only) (type yes if so):
Rename = 'yes'

#-----
#          SEARCH OF IMAGES AND IMPORT IN PROJECT
#-----

index = []

def Import_images(dirpath):

    for dirpath, dirname, filenames in os.walk(dirpath):
        # Parcourt des fichiers
        for tif_file in filenames:
            # Recherche du suffixe (.tif)
            if tif_file.endswith(Suffixe) and tif_file:
                index.append(tif_file)
            # Chemin complet de l'image
            full_path = os.path.join(dirpath, tif_file)

            grass.message('Fichier : %s' % (tif_file))
            grass.message('Dossier source : %s' % (dirpath))

    # Appel du module d'import
    grass.run_command('r.in.gdal',
```

```

        flags = 'o',
        input = full_path,
        output = tif_file,
        quiet = True,
        overwrite = True)
liste_images = index
file_list=[]
for liste in liste_images:
    file_list.append(str('%s\n' %(liste)))
date_import = str(datetime.datetime.now()).replace(':', '-')
file_txt=open('%s\DATA_LIST_%s.txt' %(output, date_import[:19]),'w+')
file_txt.writelines(file_list)
file_txt.close()

Import_images(input)

#-----
#          ATMOSPHERIC CORRECTIONS
#-----

elevation_km_neg = (Elevation_moyenne-(Elevation_moyenne*2)/1000)

#Bands numbers:
# 5 bands (R, G, B, NIR, SWIR)
S2A = '169', '168', '167', '173', '177'
S2B = '182', '181', '180', '186', '190'
# 13 bands (normal order)
S2A_all_bands = 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178
S2B_all_bands = 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191

#Function of ATCORR module, called in the Correction() function.
def Atcorr(inp, out, param):
    grass.run_command(i.atcorr
    (overwrite = True,
    flags = 'r',
    input = inp,
    output = out,
    parameters = param,
    range = [1,10000],
    rescale = [1,10000]))

#Creation of the 6S parameters and calling of the atmospheric correction function:
def Correction():
    for image in index:

        #Defining of GRASS region according to the scene:
        grass.run_command('g.region', raster= '%s.1' % (image), zoom= '%s.1' % (image))

        #Extraction of scene geographical center:
        centre_region = grass.parse_command('g.region', raster= '%s.1' % (image), flags='bg')
        centre = []
        for c in centre_region.iteritems():
            if "ll_clat" in c:
                centre.append(float([c[-1].replace("u", ""))])
            if "ll_clon" in c:
                centre.append(float([c[-1].replace("u", ""))])
        sat = image[2]

        #Extraction of acquisition time:
        date = image[:25]

        if date[15] != '0':
            mois = date[15:17]
        if date[15] == '0':
            mois = date[16]
        if date[17] != '0':
            jour = date[17:19]
        if date[17] == '0':
            jour = date[18]

```

```

if date[20] != '0':
    heure = date[20:22]
if date[20] == '0':
    heure = date[21]
if date[22] != '0':
    minutes = int(round(float(date[22:24])/60*100))
if date[22] == '0':
    minutes = int(round(float(date[23])/60*100))

for num in range(Nombre_de_bandes + 1):
    if num != 0:
        grass.run_command('r.null',
            map = '%s.%s' %(image, num),
            setnull = 0)
    if num != 0:
        param = []
        if sat == 'A':
            param.append(str('25\n'))
        if sat == 'B':
            param.append(str('26\n'))
        param.append(str('%s %s %s.%s %s %f %s %f\n' % (mois, jour, heure, minutes, centre[0], centre[1])))
        param.append(str('%s\n' %(Modele_atmospherique)))
        param.append(str('%s\n' %(Modele_aerosols)))
        param.append(str('%s\n' %(Visibilite)))
        param.append(str('%s\n' %(elevation_km_neg)))
        param.append(str('-1000\n'))
        #Selection of the bands, 5 or 13:
        if sat == 'A' and Nombre_de_bandes == 5:
            param.append(str('%s' %(S2A[num-1])))
        if sat == 'A' and Nombre_de_bandes == 13:
            param.append(str('%s' %(S2A_all_bands[num-1])))
        if sat == 'B' and Nombre_de_bandes == 5:
            param.append(str('%s' %(S2B[num-1])))
        if sat == 'B' and Nombre_de_bandes == 13:
            param.append(str('%s' %(S2B_all_bands[num-1])))
        file=open(os.path.join(input, 'params.6S.%s.%s.txt' % (image, num)), 'w+')
        file.writelines(param)
        file.close()
        prm=os.path.join(input, 'params.6S.%s.%s.txt' % (image, num))

#The ATCORR module is used here. If an error is identified, it ignores it and continues.
#Even more attention to the resulting rasters is therefore required.
try:
    Atcorr('%s.%s' %(image, num), '%s.atcorr.%s' %(image, num), prm)
except:
    pass
if Suppr == 'Oui' or Suppr == 'OUI' or Suppr == 'oui' or Suppr == 'YES' or Suppr == 'Yes' or Suppr == 'yes':
    os.remove('%s' % (prm))

-----
EXPORT OF CORRECTED IMAGES
-----

Group():
for image in index:
    grass.run_command('g.region', raster= '%s.1' % (image), zoom= '%s.1' % (image))
    grass.run_command('i.group',
        overwrite = True,
        group = '%s.corr' % (image),
        input = ('%s.atcorr.1' % (image), '%s.atcorr.2' % (image), '%s.atcorr.3' % (image),
            '%s.atcorr.4' % (image), '%s.atcorr.5' % (image)))

In this export function, the flags -c and -m are used here to skip the export of the color table (-c)
and to not write non-standart metadata (-m), in the objective of enhancing compatibility with other
GIS softwares. The flags can be removed or changed by removing the line or adding or removing flags.
Export():
for image in index:
    name = []

```

```

if Rename == 'yes' or Rename == 'Yes' or Rename == 'oui' or Rename == 'Oui':
    print'sup'
    name = ('%s-%s-%s-%s-%s-%s_S2%s%s' % (image[11:15], image[15:17], image[17:19],
    image[20:22], image[22:24], image[24:26], image[2], Suffixe))
if not Rename or Rename == "":
    name = ('corr_%s' % (image))
#Declaration de la region selon l'image selectionnee:
grass.run_command('g.region', raster= '%s.atcorr.1' %(image), zoom= '%s.atcorr.1' %(image))
grass.run_command('r.out.gdal',

                                                                    overwrite = True,

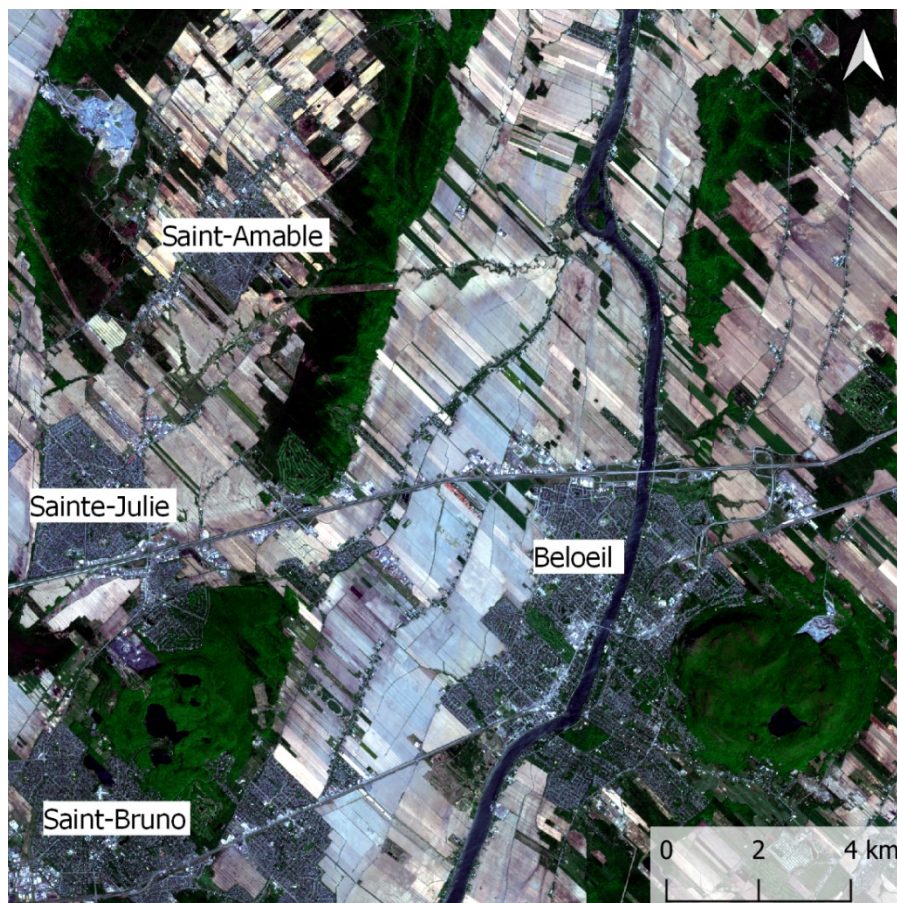
                                                                    input = '%s.corr' % (image),

    flags = 'cm',

    format = 'GTiff',
    output = '%s\\%s' % (output, name))

# Calling of the correction, grouping and export functions.
# If desired, put in commentary to skip.
Correction()
Group()
Export()
print('Process is finished')

```



Comme on peut le voir sur cette image prise au printemps 2018, bien que la végétation aie reprise le dessus dans les zones densément couvertes, les sols agricoles sont encore nus. Leur texture est donc analysable à l'aide de leur réflectance.

Figure 14. Agrandissement sur l'image Sentinel-2 #1 (voir Tableau 2, section [3.2.1.](#))





Figure 15. Granulométrie de la zone d'étude, classifiée avec l'image #11, du 2018/06/10



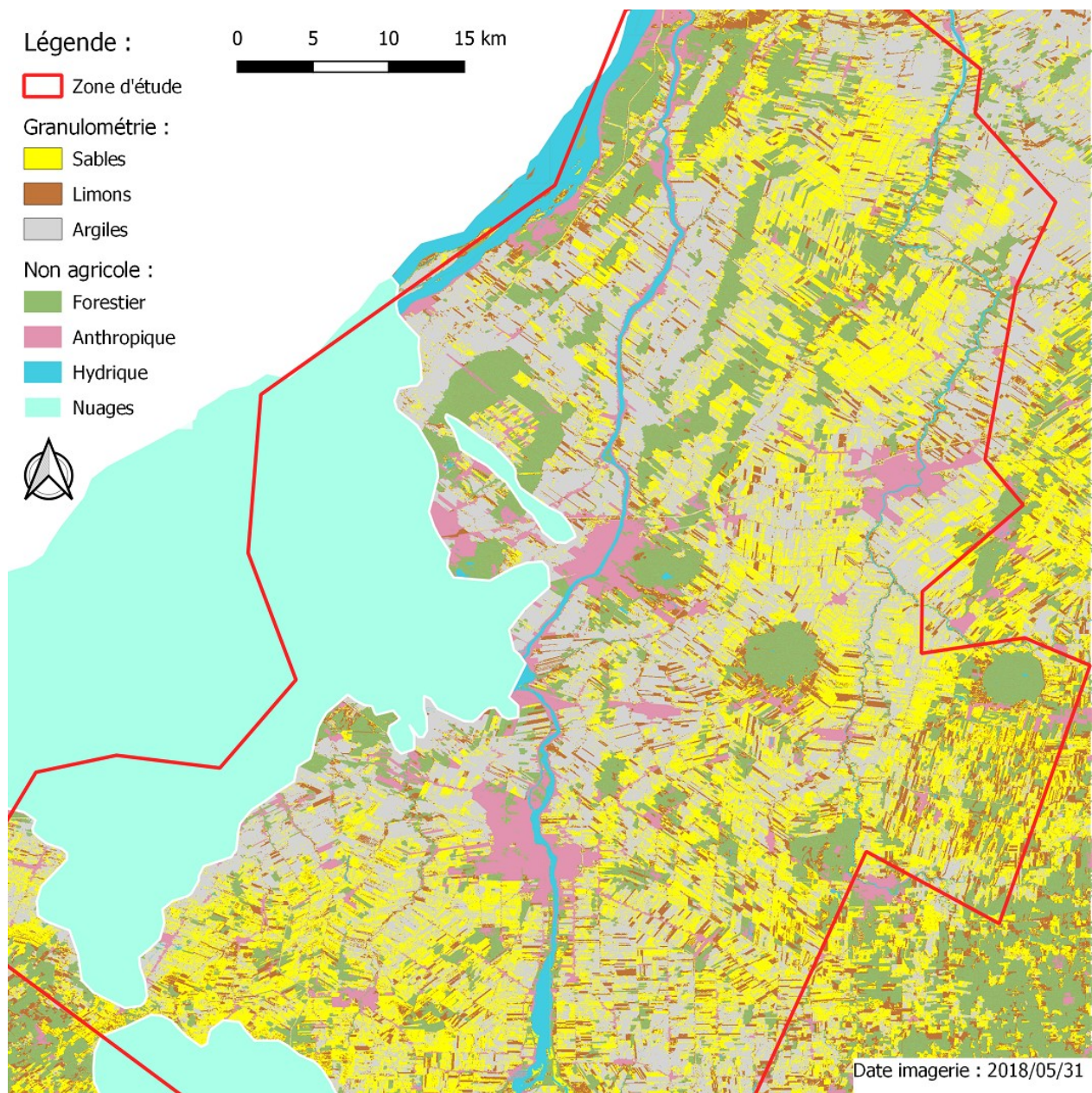


Figure 16. Granulométrie de la zone d'étude, classifiée avec l'image #10, du 2018/05/31



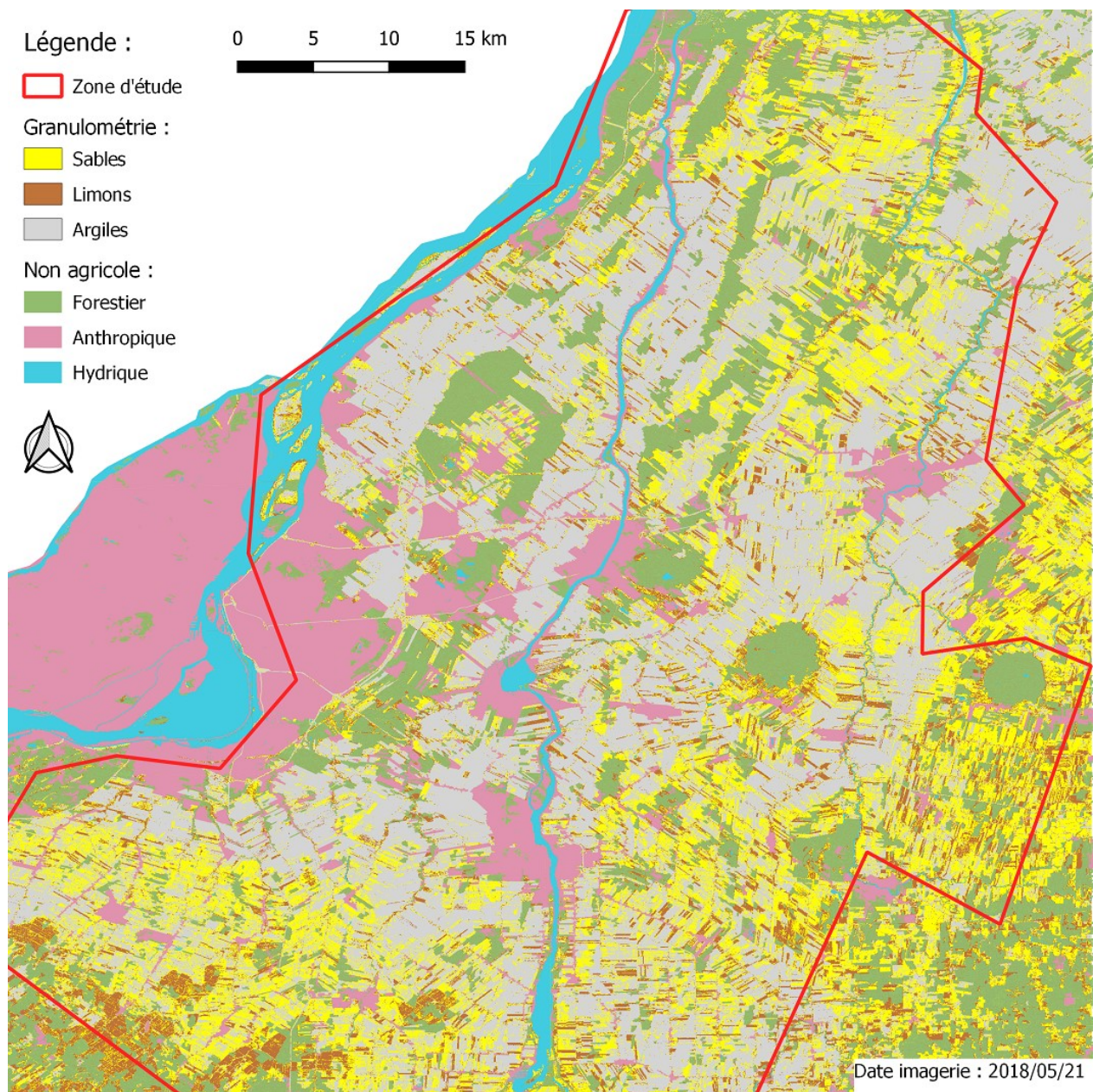


Figure 17. Granulométrie de la zone d'étude, classifiée avec l'image #9, du 2018/05/21



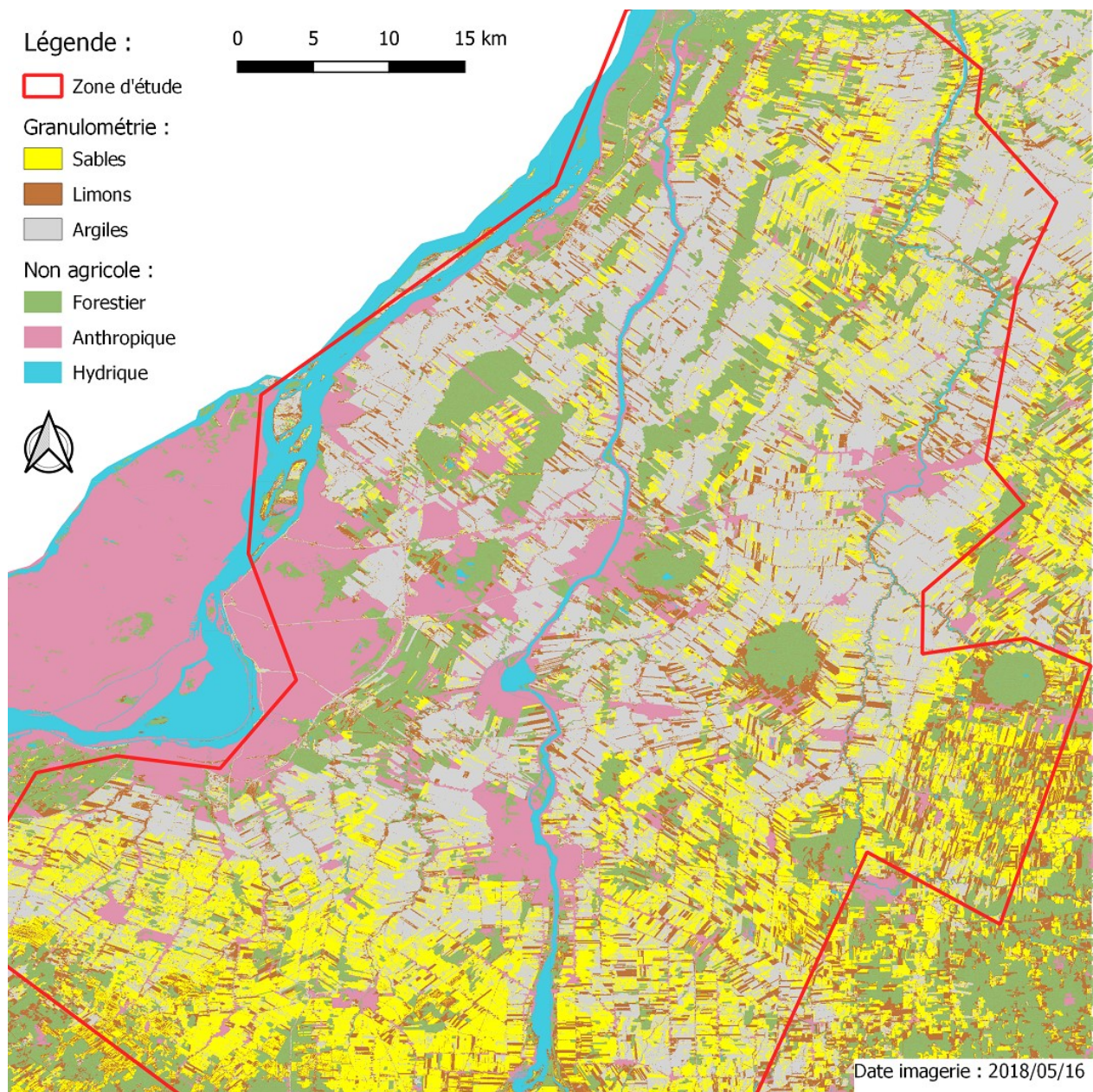


Figure 18. Granulométrie de la zone d'étude, classifiée avec l'image #8, du 2018/05/16



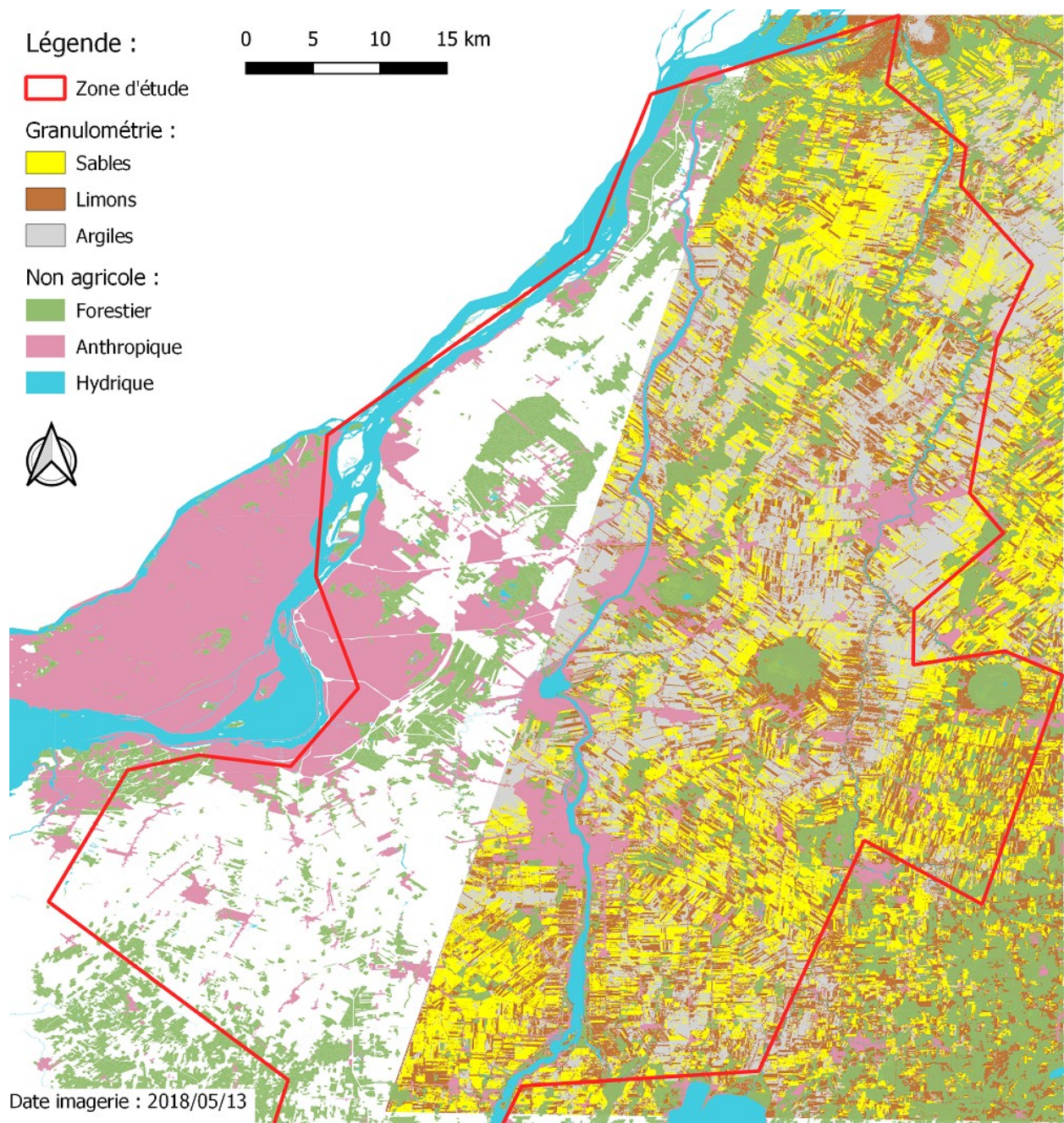


Figure 19. Granulométrie de la zone d'étude, classifiée avec les images #6 et 7, du 2018/05/13



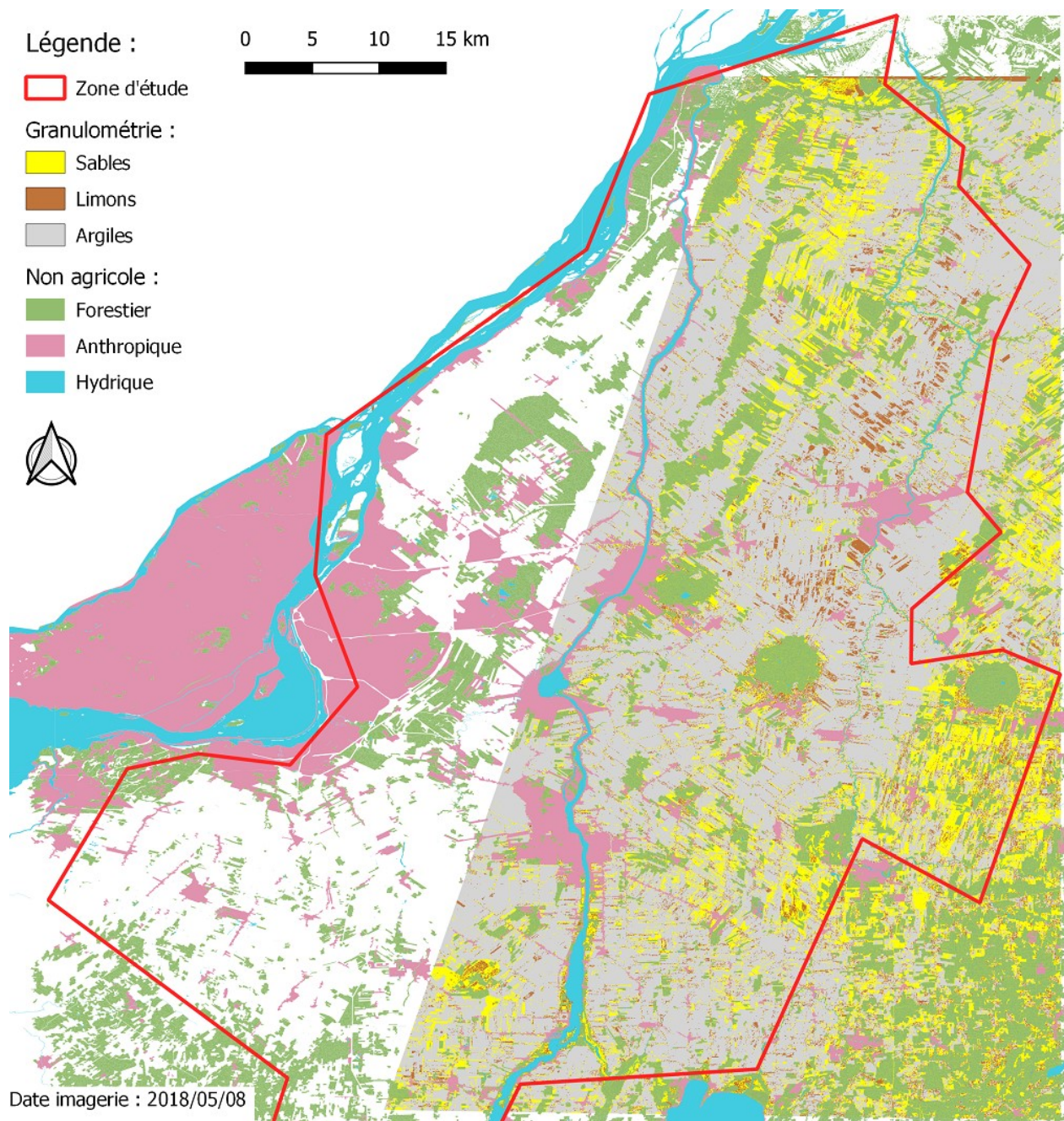


Figure 20. Granulométrie de la zone d'étude, classifiée avec l'image #5, du 2018/05/08



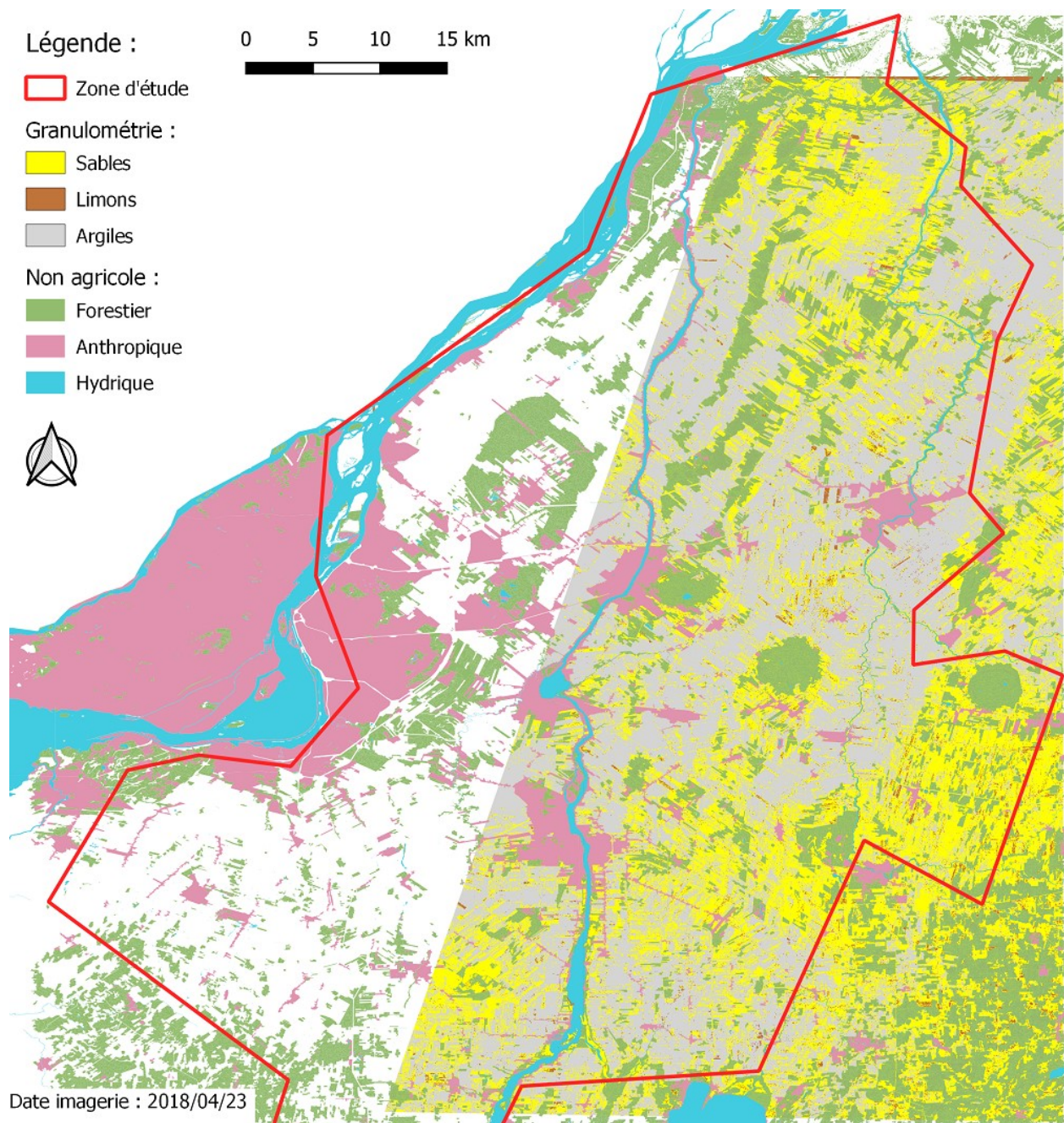


Figure 21. Granulométrie de la zone d'étude, classifiée avec l'image #4, du 2018/04/23

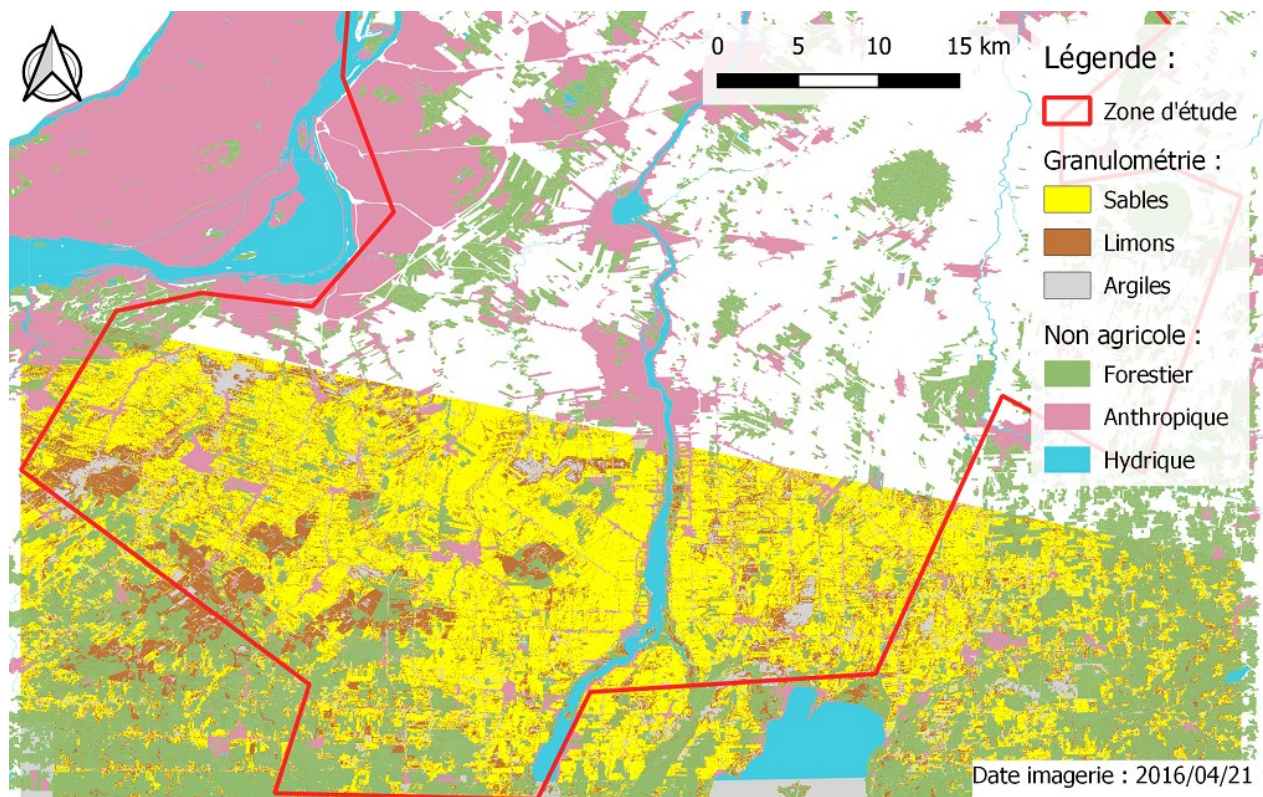


Figure 22. Granulométrie de la zone d'étude, classifiée avec les images #1,2 et 3, du 2016/04/21



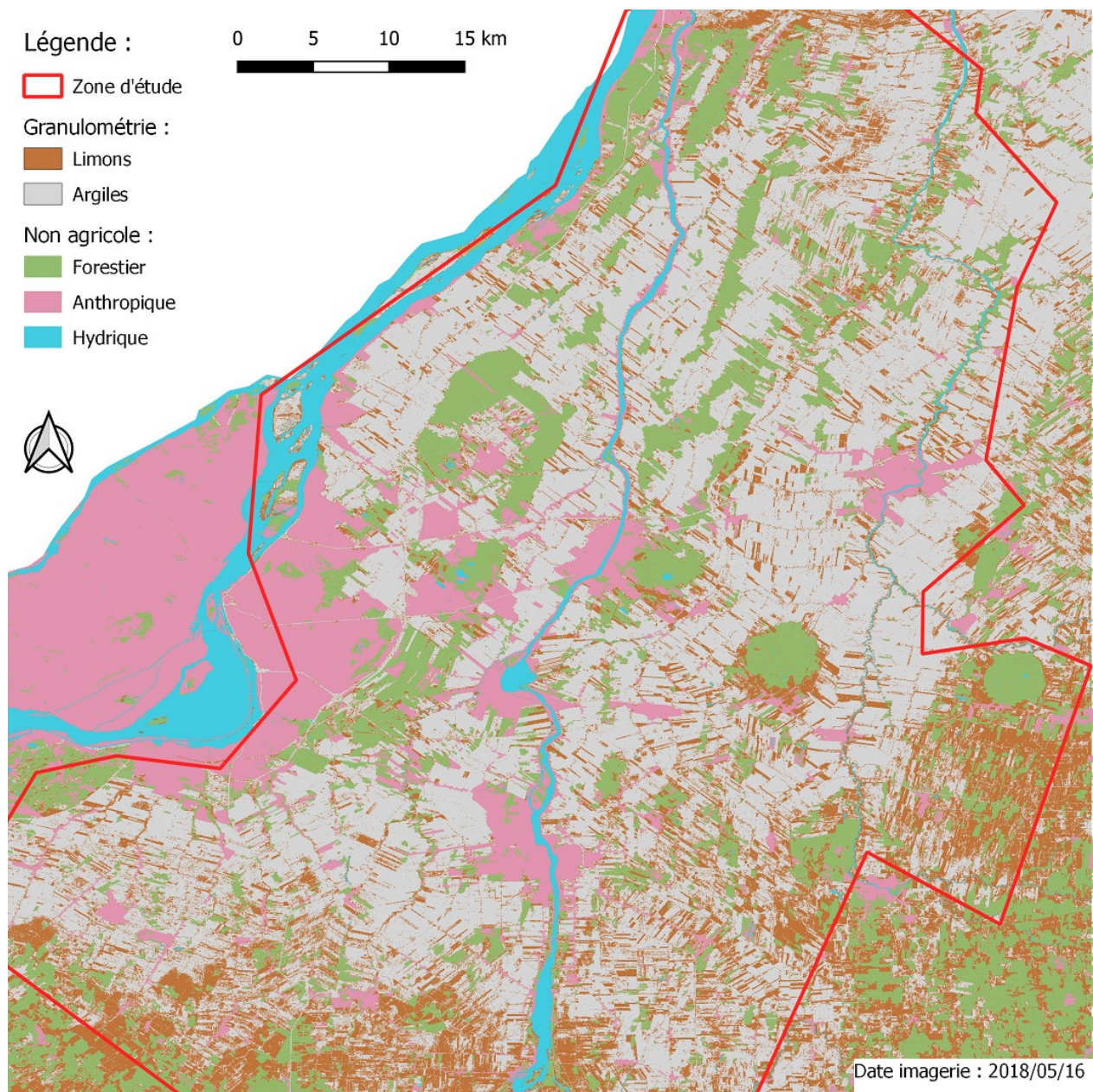


Figure 23. Granulométrie de la zone d'étude, classifiée avec l'image #8, du 2018/05/16 (2 classes)



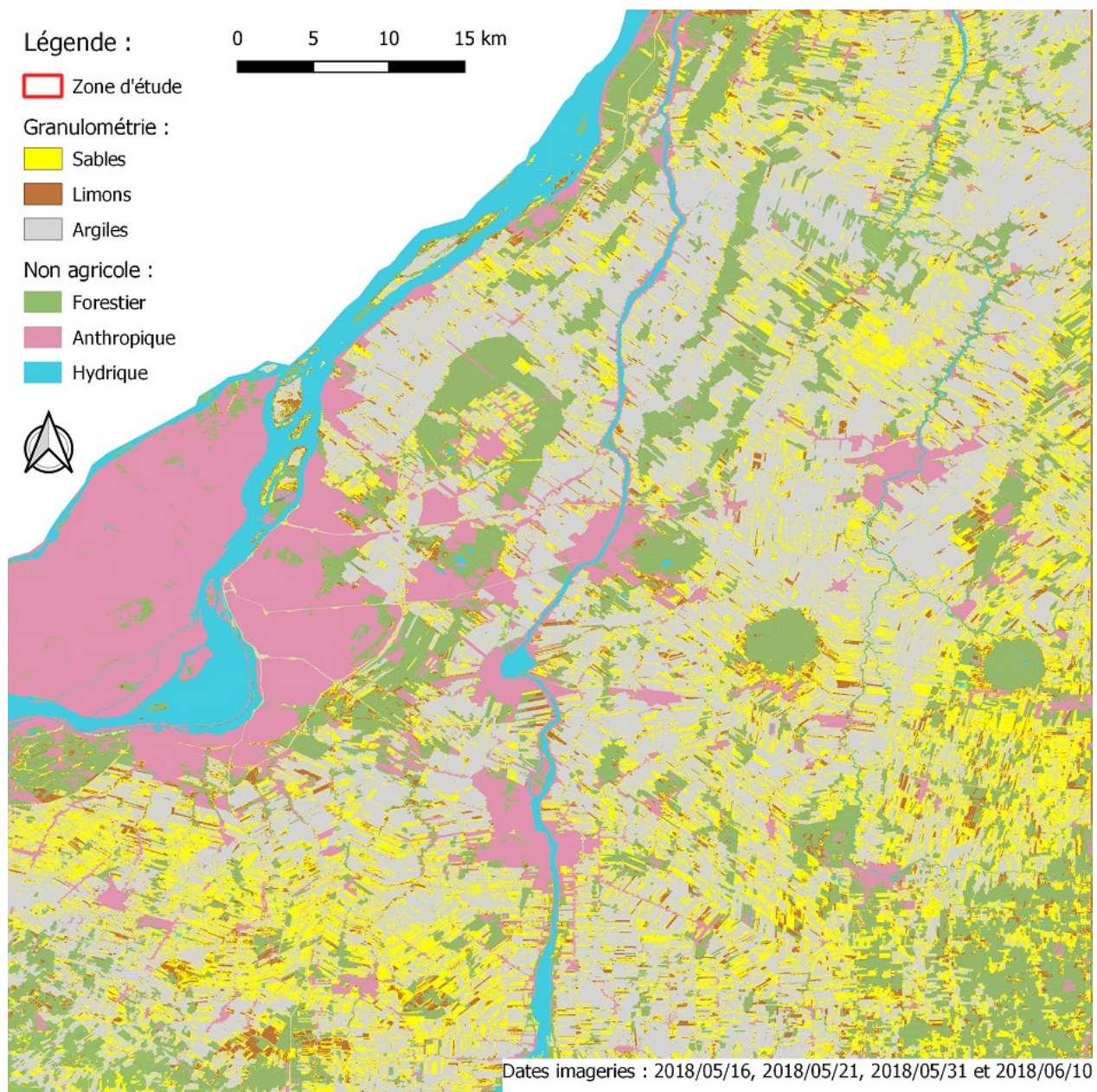


Figure 24. Granulométrie de la zone d'étude, classifiée avec les couvertures du 18/05/16, 18/05/21, 18/05/13 et 18/06/10



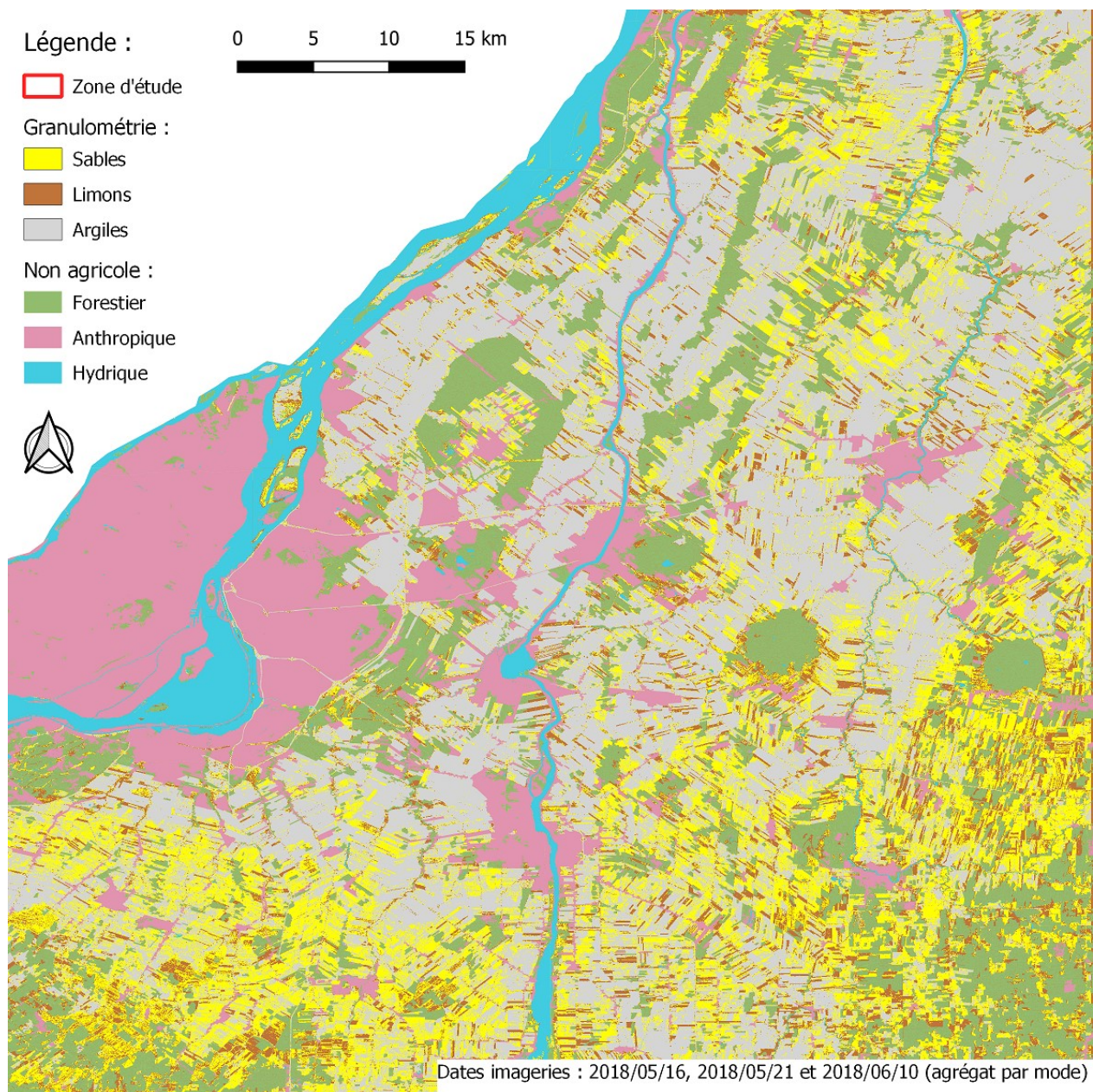


Figure 25. Granulométrie de la zone d'étude, calculée par agrégat (par mode) post-classification des couvertures du 18/05/16, 18/05/13 et 18/06/10





On peut remarquer que le sol, bien que légèrement végétalisé par de jeunes pousses par endroits, est encore relativement nu. Il s'agit d'une photo d'une culture de maïs qui a été prise à proximité de Saint-Basile-Le-Grand, dans la zone d'étude

Figure 26. Photo de proximité d'une terre agricole, 16 mai 2019